Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/NZ04/000333

International filing date: 22 December 2004 (22.12.2004)

Document type: Certified copy of priority document

Document details: Country/Office: NZ
Number: 530331
Filing date: 22 December 2003 (22.12.2003)

Date of receipt at the International Bureau: 17 January 2005 (17.01.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

Ministry of Economic Development
Manatū Ōhanga
Intellectual Property Office of New Zealand

PCT/NZ2004/000333

CERTIFICATE

This certificate is issued in support of an application for Patent registration in a country outside New Zealand pursuant to the Patents Act 1953 and the Regulations thereunder.

I hereby certify that annexed is a true copy of the Provisional Specification as filed on 22 December 2003 with an application for Letters Patent number 530331 made by AgResearch Limited.

Dated 11 January 2005.

Neville Harris
Commissioner of Patents, Trade Marks and Designs

COMMISSIONER OF PATENTS, TRADEMARKS & DESIGNS
NEW ZEALAND

330 High Street, Lower Hutt
PO Box 30687, Lower Hutt, New Zealand or DX SX 11129, Wellington
Freephone: 0508 4 IPONZ (0508 447 669), International callers dial: +64 3 962 2607, Fax: +64 4 560 1691, www.iponz.govt.nz

530331

PATENTS FORM NO. 4

Appln Fee: $50.00

James & Wells ref: 42616/29

PATENTS ACT 1953

PROVISIONAL SPECIFICATION

INDOLE-DITERPENE BIOSYNTHESIS

We, AgResearch Limited, a New Zealand company of East Street, Ruakura Campus, Hamilton, New Zealand do hereby declare this invention to be described in the following statement:

INDOLE-DITERPENE BIOSYNTHESIS

TECHNICAL FIELD

The present invention relates to the biosynthesis of indole diterpene compounds. In particular, the invention relates to genes encoding enzymes considered responsible for the synthesis of lolitrems. A further aspect of this invention is the use of these genes to modify or transfer the lolitrem biosynthetic pathway in endophytes. In a further aspect of this invention these genes can be modified and expressed in transgenic plants to increase resistance to insects. An even further aspect of this invention is the modification of the lolitrem biosynthetic pathway to produce novel intermediates that may have biomedical applications.

BACKGROUND ART

Indole-Diterpenes

The indole-diterpenes are a large, structurally diverse group of natural products principally found in filamentous fungi notably of the genera *Penicillium*, *Aspergillus*, *Claviceps*, Epichloeand *Neotyphodium* (Steyn and Vleggaar 1985; Mantle 1987; Scott et al. 2003). They may be classified into the following structural sub-groups, the penitrems, janthitrems, sulphinines (Laakso et al., 1992), nodulisporic acid (Ondeyka et al., 1997) and thiersinines (Li et al., 2002). These metabolites all have a common core structure comprised of a cyclic diterpene skeleton derived from geranylgeranyl diphosphate (GGPP) and an indole moiety derived from either tryptophan or a tryptophan precursor (Acklin et al. 1977; de Jesus et al. 1983; Laws and Mantle 1989). Further complexity of

the carbon skeleton is achieved by additional prenylations, different patterns of ring substitutions and different ring stereochemistry. Many of these compounds are potent mammalian tremorgens (Cole and Cox 1981) while others are known to have confer anti-insect activity (Gloer 1995).

Paxilline Biosynthesis

Until recently, very little was known about the pathways for the biosynthesis of the indole-diterpenes, although putative biosynthetic schemes have been proposed on the basis of chemical identification of likely intermediates from the organism of interest and related filamentous fungi (Mantle and Weedon 1994; Munday-Finch et al. 1996; Gatenby et al. 1999). The recent cloning and characterization of a cluster of genes from *Penicillium paxilli* required for the biosynthesis of paxilline has provided for the first time an insight into the genetics and biochemistry of indole-diterpene biosynthesis (Young et al. 2001).

Key genes identified in this cluster include a GGPP synthase (*paxG*), a FAD-dependent monooxygenase (*paxM*), a prenyl transferase (*paxC*) and two cytochrome P450 monooxygenases, *paxP* and *paxQ*. Deletion of *paxG* resulted in mutants that were paxilline negative, confirming that this gene is essential for paxilline biosynthesis (Young et al. 2001). Targeted deletion of *paxM* and *paxC* in *P. paxilli* also result in mutants that are defective in paxilline biosynthesis (B. Scott, L. McMillan, J. Astin, C. Young, E. Parker, unpublished results). It is proposed that PaxM and PaxC are required to catalyse the addition of indole-3-glycerol phosphate to GGPP and subsequent cyclisation to form the first stable indole-diterpene, possibly paspaline (Parker and Scott 2004). Deletion of *paxP* and *paxQ* give rise to strains that accumulate paspaline and 13-desoxypaxilline, respectively, suggesting that these are the substrates for the corresponding enzymes

(McMillan et al. 2003). Overall, these results establish that at least 5 genes are required for the biosynthesis of paxilline in *P. paxilli*.

The identification of a geranyl-geranyl diphosphate (GGPP) synthase gene (*paxG*) within this cluster, and confirmation by deletion analysis that it is necessary for paxilline biosynthesis, suggest that the synthesis of GGPP is one of the first steps in the synthesis of this indole-diterpene (Young *et al.* 2001). *P. paxilli*, like *Gibberella fujikuroi* (Mende *et al.* 1997: Tudzynski and Hölter 1998), recently renamed *Fusarium fujikuroi* (O'Donnell *et al.* 1998), has two GGPP synthase genes, but the second, *ggsI*, is unable to complement the *paxG* deletion, presumably because of cellular partitioning of the two enzymes (Young *et al.* 2001). The synthesis of paxilline is predicted to involve several oxygenation steps (Munday-Finch *et al.* 1996), and the presence within the cluster of genes for two FAD-dependent monooxygenases (*paxM* and *paxN*) and for two cytochrome P450 monooxygenases (*paxP* and *paxQ*) is consistent with this chemistry (Young *et al.* 2001).

The only other fungal diterpene gene cluster reported to date is that for the biosynthesis of gibberellins in *Fusarium fujikuroi* (teleomorph *Gibberella fujikuroi*)(Tudzynski and Hölter 1998). This cluster also includes a GGPP synthase gene, *ggs-2*, required for the first committed step in gibberellin biosynthesis. Interestingly, both fungal species contain an additional copy of a GGPP synthase gene, *ggs1* in *P. paxilli* (Young et al. 2001) and *ggs-1* in *F. fujikuroi* (Mende et al. 1997). This suggests that the presence of two copies of GGPP synthases could be a molecular signature for diterpene biosynthesis in filamentous fungi, one copy being required for primary metabolism and the second for secondary (diterpene) metabolism. Given that genes for secondary metabolite biosynthesis in fungi are generally organised in clusters (Keller and Hohn 1997), molecular cloning of GGPP synthases combined with chromosome walking provides a rapid strategy for cloning new indole-diterpene gene clusters.

Lolitrems

Epichloë/Neotyphodium endophytes are a group of Clavicipitaceous fungi (Clavicipitaceae, Ascomycota) that form symbiotic associations with temperate grasses such as perennial ryegrass and tall fescue (Schardl 2001; Scott 2001). The plant provides nutrients for the endophyte and a means of dissemination through the seeed. The endophyte protects the host from biotic (e.g. insect and mammalian herbivory) and abiotic stress (e.g. drought). Fungal synthesis of secondary metabolites appears to be the main mechanism for protection of the symbiotum from herbivory.

The ability of *Epichloë/Neotyphodium* endophytes to synthesize bioprotective metabolites *in planta* constitutes a major ecological benefit for the symbiotum (Schardl 1996). Metabolites identified to date include both anti-insect (e.g. peramine and lolines) and anti-mammalian (ergot alkaloids and indole-diterpenes)(Bush et al. 1997). However, from an agricultural perspective endophyte production of mammalian toxins such as the indole-diterpenes, and in particular lolitrem B, is detrimental to grazing livestock. Consequently, there is considerable commercial interest in developing associations containing endophytes that are not toxic to mammals (Fletcher 1999; Popay et al. 1999).

The lolitrems are produced by the *Epichloë* endophytes in association with temperate grass species (Gallagher et al., 1984). These fungi are often found as an infection in perennial ryegrass (*Lolium perenne*) and tall fescue grasses (*Festuca arundinacea*).

Endophytes are symbiotic fungi and are prevalent in New Zealand pastures. The fungal metabolites from these endophytes are thought to serve as chemical defence systems for the fungi that produce them. They may also be of use in protecting the food source from consumption by other organisms (US 4,973,601).

However of these fungi also pose a problem in that, at least lolitrem B, is known to be the main causative agent in ryegrass staggers (Fletcher and Harvey, 1981). This is a condition in which animals grazing on endophyte infected pastures develop ataxia, tremors, and hypersensitivity to external stimuli. The lolitrem neurotoxin (staggers) reaction is long acting but is however completely reversible (Smith et al 1997, McLeay et al 1999). The time course of tremors induced by lolitrem B is dramatically different from that of other indole diterpenes, for example paxilline and analogues. Paxilline analogues induce tremors of rapid onset and short duration while tremors induced by lolitrem derivatives take hours to reach maximum intensity and last for days.

The mechanism by which lolitrem B and related indole-diterpenes cause tremorgenicity in mammals is not well defined but biochemical and clinical studies indicate that these effects are due in part to effects on receptors and interference with neurotransmitter release in the central and peripheral nervous system (Selala et al. 1991). Some have been shown to potentiate chloride currents through $GABA_A$ receptor chloride channels heterologously expressed in *Xenopus* oocytes (Yao et al. 1989). Many of the indole-diterpenes are potent inhibitors of high conductance Ca^{2+}- activated K^+ (maxi-K) channels (Knaus et al. 1994; McMillan et al. 2003)

All references, including any patents or patent applications cited in this specification are hereby incorporated by reference. No admission is made that any reference constitutes prior art. The discussion of the references states what their authors assert, and the applicants reserve the right to challenge the accuracy and pertinency of the cited documents. It will be clearly understood that, although a number of prior art publications are referred to herein, this reference does not constitute an admission that any of these documents form part of the common general knowledge in the art, in New Zealand or in any other country.

It is acknowledged that the term 'comprise' may, under varying jurisdictions, be attributed with either an exclusive or an inclusive meaning. For the purpose of this specification, and unless otherwise noted, the term 'comprise' shall have an inclusive meaning - i.e. that it will be taken to mean an inclusion of not only the listed components it directly references, but also other non-specified components or elements. This rationale will also be used when the term 'comprised' or 'comprising' is used in relation to one or more steps in a method or process.

It is an object of the present invention to address the foregoing problems or at least to provide the public with a useful choice.

Further aspects and advantages of the present invention will become apparent from the ensuing description which is given by way of example only.

DISCLOSURE OF INVENTION

For the purposes of this specification the terms 'pax' and '*ltm*' will be used interchangeably in relation to the gene names. It should be appreciated by those skilled in the art that the specific genes are the same, for example paxC and *ltmC* however for naming conventions, the prefix used relates to the compound expressed by the gene i.e. paxilline in the case of pax and lolitrem in the case of *ltm*.

Further, for the purposes of the specification, the term 'biosynthesis' refers to the building up of a chemical compound in the physiological processes of a living organism.

According to one aspect of the present invention there is provided isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked

to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, which encode enzymes that catalyse at least one route in the biosynthesis of the indole diterpene.

Preferably, the isolated nucleic acid molecules are selected from *ltmG*, *ltmM*, *ltmK*, *ltmP*, *ltmQ*, *ltmD*, *ltmJ*, and combinations thereof.

More preferably, the isolated nucleic acid molecules are selected from *ltmG*, *ltmM*, *ltmK*, *ltmP*.

Most preferably, the isolated nucleic acid molecules, *ltmG*, *ltmM*, *ltmK,* form a tightly linked cluster.

Preferably, the isolated nucleic acid, substantially as described above includes sequences as described hereafter.

According to a further aspect of the present invention there is provided the use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, to identify further lolitrem biosynthetic genes.

Preferably, the isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, encodes enzymes that synthesise indole diterpenes.

Preferably, indole diterpenes are lolitrem compounds.

Preferably, identifying is completed by linkage on a genetic or physical map, or by further homology screening.

According to a further aspect of the present invention there is provided the use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid

sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster in the biosynthesis of indole diterpene compounds.

According to a further aspect of the present invention there is provided a method of biosynthesising indole diterpene compounds, which method includes use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster.

Preferably, the nucleic acids encode enzymes that synthesise indole diterpenes.

Preferably, indole diterpenes are lolitrem compounds.

According to a further aspect of the present invention there is provided use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster to genetically characterise a candidate lolitrem biosynthetic gene cluster or individual lolitrem biosynthetic genes.

Preferably, such characterisation includes but is not limited to gene expression analysis, targeted gene disruption, expression in heterologous expression systems. It should be appreciated by those skilled in the art that other methods of characterisation may also be used without departing from the scope of the invention.

According to a further aspect of the present invention there is provided a method of screening for the properties of a fungal species by identification of biosynthesis nucleic acids present in the fungi and comparing this to known biosynthesis pathways to determine compounds that the fungi may produce.

In preferred embodiments, the use or method as described above is used to provide an endophyte which does not produce detectable levels of lolitrems and isot observably

tremorgenic. The said endophyte may be derived from the *Epichloë* or *Neotyphodium* genus and may be naturally occurring or be generated by transgenesis including mutagenesis or silencing of the lolitrem biosynthetic pathway by gene disruption using homologous recombination, by gene silencing using antisense or RNAi technologies.

According to a further aspect of the present invention, there is provided the use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, to screen for endophyte strains for alterations or mutations in the lolitrem biosynthetic pathway.

Preferably, these include non-tremorgenic strains/isolates or strains with increased insecticidal activity including those that produce lolitrem intermediates and/or shearinines and/or janthitrems. In preferred embodiments, this screening includes but is not limited to methods including, southern blot analysis, PCR, SNP analysis, Microsatellite or SSR analysis, sequencing, real-time PCR or any other molecular techniques.

According to a further aspect of the present invention there is provided a method of producing indole diterpene compounds or their intermediate compounds by use of isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster.

Preferably, the isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, encode enzymes that synthesise indole diterpenes.

Preferably, the indole diterpene compounds are lolitrem compounds. Preferably, compounds expressed are controlled by manipulation of the genes present.

Preferably, the pathway achieved is completed in conjunction with an expression system.

Preferably, the expression of the isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster is in a heterologous expression system (for example bacteria, yeast, fungi, plants, animal cells) to produce indole diterpenes or their intermediates

Preferably, the expression system includes a plant or fungi. Most preferably, the expression system is an endophytic fungus. In preferred embodiments, the said endophyte may be derived from the *Epichloë* or *Neotyphodium* genus.

According to a further aspect of the present invention there is provided manipulated isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, which produce an effect or effects selected from: a less toxic effect, a more toxic effect, a desired agricultural effect, a desired biochemical effect, a desired neurological effect, a desired insecticidal effect, and combinations thereof.

In preferred embodiments, the toxic effect is controlled by manipulation of the gene expression to alter the level of lolitrems produced. Most preferably, the expression is manipulated to control the degree of tremorgenicity produced in animals consuming lolitrem containing herbage.

According to a further aspect of the present invention there is provided a method of producing a transgenic plant by expressing isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, into the plant.

In preferred embodiments, the isolated nucleic acid molecules, sequences or fragments, or any other nucleic acid sequence physically linked to within 250 kilobase pairs of the lolitrem biosynthetic gene cluster, encode enzymes that synthesise indole diterpenes.

Preferably, indole diterpenes are lolitrem compounds.

According to a further aspect of the present invention there is provided the alteration of the lolitrem biosynthetic pathway by transgenesis to produce insecticidal indole diterpenes. Preferably these include but are not limited to: lolitrem intermediates and/or shearinines, and/or janthitrems.

Preferably alteration of the lolitrem biosynthetic pathway is completed to produce a plant so with insecticidal properties.

According to a further aspect of the present invention there is provided seeds of a transgenic plant containing biosynthetic genes for production indole diterpenes.

It should be appreciated from the above description that there is provided nucleic acid molecules for the biosynthesis of indole diterpene compounds. It will be appreciated further that through knowledge of these molecules, further molecules can be determined that relate to aspects of the biosynthesis process. Further, it will be appreciated that the genes have a variety of resulting applications such as screening to determine biosynthesis products and manipulation of the genes that influence the biosynthesis to create desired intermediate and end product indole diterpene compounds.

BRIEF DESCRIPTION OF DRAWINGS

Further aspects of the present invention will become apparent from the following description which is given by way of example only and with reference to the

accompanying drawings in which:

Figure 1. Structure of lolitrem B

Figure 2. Degenerate PCR and Southern hybridisation of the GGPP synthase gene fragments CY28 and CY29 Degenerate PCR analysis using primers (**A**) g27 and g28, and (**B**) g27 and g29. Lane (1) 1 kb+ ladder, (2) *N. lolii* strain Lp19, (3) *E. festucae* strain FI1, (4) E. *typhina* strain E8 (5) wild-type *P. paxilli*, (6) *P. paxilli* strain LM662, (7) blank. Southern hybridisation of the ggs fragments. (**C**) probed with fragment CY29 (*ggs1*). (**D**) probed with frgament CY28 (*ltmG*). Lane (1, 4 and 7) *N. lolii* strain Lp19, (2, 5 and 8) *E. festucae* strain FI1 (3, 6 and 9) *E. typhina* strain E8. Lanes 1-3 are digested with *Eco*RI, lanes 4-6 are digested with *Hin*dIII and lanes 7-9 are digested with *Sst*I. The size standards are in kb.

Figure 3 . *N. lolii* and *E. festucae* lolitrem gene cluster. Physical map of the (**A**) Lp19 and (**B**) FI1 lolitrem gene cluster. The CY28 PCR fragment used as a probe to isolate lambda clones, is a green box. Each gene is shown as a black rectangle with intron marked and an arrow above the genes shows the gene direction. The yellow box is a microsatellite with a core sequence of TAATG. The red and blue boxes are the fragments used to make the *ltmM* knockout construct. The retrotransposons, Tahi and Rua, are shown as red and blue lines with arrow heads as the LTR sequences. Each fragment used as a probe is a green oval placed under the region of the probe. (**C**) The *ltmM* knockout construct, pCY39. (**D**) The PCR screen for a knockout in FI1. Lanes (1) 1kb+ ladder, (2) CYFI1M-28, (3) CYFI1M-142, (4) CYFI1M-61, (5) CYFI1M-151, (6) FI1, (7) pCY39, (8) H_2O control. The 7-kb *Xho*I fragment used for preparing the complementation construct is also shown.

Figure 4. The nucleotide sequence of *N. lolii* strain Lp19 *ltmG*.

Figure 5. The polypeptide sequence of *N. lolii* strain Lp19 *LtmG*.

Figure 6. The nucleotide sequence of *N. lolii* strain Lp19 *ltmM*.

Figure 7. The polypeptide sequence of *N. lolii* strain Lp19 *LtmM*.

Figure 8. The nucleotide sequence of *N. lolii* strain Lp19 *ltmK*.

Figure 9. The polypeptide sequence of *N. lolii* strain Lp19 *LtmK*.

Figure 10. The nucleotide sequence of *N. lolii* strain Lp19 *ltmG*, *ltmM* and *ltmK* gene cluster.

Figure 11. The nucleotide sequence of *E. festucae* strain FI1 *ltmG*.

Figure 12. The nucleotide sequence of *E. festucae* strain FI1 *ltmM*.

Figure 13. The nucleotide sequence of *E. festucae* strain FI1 *ltmK*.

Figure 14. The polypeptide sequence of *E. festucae* strain FI1 *LtmG*

Figure 15. The polypeptide sequence of *E. festucae* strain FI1 *LtmM*

Figure 16. The polypeptide sequence of *E. festucae* strain FI1 *LtmK*

Figure 17. HPLC analysis for lolitrem B production in endophyte infected ryegrass. The plant tissue was harvested mid summer.

Figure 18. Structure of paspaline.

Figure 19. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxP.

Figure 20. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxP

Figure 21. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxP

Figure 22. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxD

Figure 23. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxD

Figure 24. An EST derived nucleic acid fragment from the suppressive subtractive hybridization library with homology to *Penicillium paxili* paxD

Figure 25. An EST derived nucleic acid fragment from the an vitro culture library with homology to cytochrome P450 monooxygenases

BEST MODES FOR CARRYING OUT THE INVENTION

Example 1. Isolation of nucleic acid fragments containg homology to GGPP synthases from *N. lolii* and *E. festucae*

Fungal strains, *E. coli* strains, plasmids and lambda clones are described in Table 1.

Table 1: Strains, plasmids, and lambda clones.

Strain	PN number	Relevant characteristics	Reference
Lp19	PN2191	*Neotyphodium lolii*	
Fl1		*Epichloë festucae*	
E8		*Epichloë typhina*	
CYFl1-M28	PN2303	*E. festucae ΔltmM::hph*	This study
CYFl1-M61	PN2301	E. festucae *ΔltmMG::hph*	This study
CYFl1-M142	PN2296	*E. festucae ΔltmM::hph*	This study

CYFI1-M151	PN2294	*E. festucae* Δ*ltmM*::*hph* ectopic integration	This study
pCB1004		Amp^R/Hyg^R	Carroll et al 1994
pCY28		209 bp *ltmG* fragment in pGEM-T, Amp^R	This study
pCY29		272 bp *ggsA* fragment in pGEM-T, Amp^R	This study
pCY39		Amp^R / Hyg^R, *ltmM* knockout construct	This study
pGEM-T		Amp^R	Promega
pGEM-T-easy		Amp^R	Promega
pPN1688	PN1688	Amp^R /Hyg^R	This study
pUC118		Amp^R	This study
□CY218		Lp19□GEM12 containing *ltmG*	This study
□CY255		Lp19□GEM12 containing *ltmK*	This study
□CY275		Lp19□GEM12 overlapping □CY255	This study
□CY100		Lp19□GEM12 containing ggsA	This study
G1114		Nui ryegrass, CYFI1-M28	This study
G1119		Nui ryegrass, CYFI1-M61	This study
G1126		Nui ryegrass, CYFI1-M142	This study
G1130		Nui ryegrass, CYFI1-M151	This study
G1137		Nui ryegrass, FI1	This study
G1138		Nui ryegrass, endophyte free	This study

All bacteria were grown in LB medium overnight at 37°C. For maintenance, the fungal cultures were grown on 2.4% potato dextrose (PD; Difco) agar plates at 22°C until suitable growth was attained. For DNA isolation, the fungal strains were grown in PD broth at 22°C for 5-12 days. The protein sequences of the available fungal GGPPS genes from *Neurospora crassa al-3*, (accession number AAC13867)(Barbato et al. 1996) *S. cerevisiae Bts1* (accession number AAA83662) *P. paxilli paxG* (accession number AF279808) (Young et al. 2001), and *Gibberella fujikuroi ggs-1* (accession number CAA65644) (Mende et al. 1997) and *ggs-2* (accession number CAA75568) (Tudzynski and Hölter 1998) were aligned (Higgins et al. 1994) to determine conserved domains that would be suitable for degenerate primer design. Primers, ggpps27, ggpps28 and ggpps29, were designed to three highly conserved regions taking in to consideration the

placement of any known introns. The sequence of these and other primers are shown in Table 2.

Table 2: Primer list

Name	sequence 5' → 3'	amplifies
CY 4	GCT TGG ATC CGA TAT TGA AGG AGC	hph/BamHI
CY 5	TTG GAT CCG GTT CCC GGT CGG CAT	hph/BamHI
ggpps 27	CAY MGI GGT CAR GGT ATG GA	dPCR
ggpps 28	TTC ATR TAG TCG TCI CKT ATY TG	dPCR
ggpps 29	AAC TTT CCY TCI GTS ARG TCY TC	dPCR
lol 1	TGG ATC ATT CGC AGA TAC	*ltmG*
lol 2	GTG TGA GAT TAA GAC GTC	LHS
lol 3	ACC GAC GCC ATT AAT GAG	*ltmG*
lol 7	ACT GGG CAT CTT CCA TAG	*ltmM*/mid
lol 14	ATT AGA GGC ACC GAA CGC	RT-PCR *ltmM*
lol 15	ATC AAG CTG GCT ATC CTC	*ltmP*
lol 17	AAA TAA TGG GCA AGG AGC	KO PstI
lol 18	TGG GAAT TTT GGA AAT GGC	KO PstI
lol 28	GCT CCT TGC CCA TTA TTT	RT-PCR *ltmM*
lol 29	GTC TTG ATC GTC TGC ATC	RT-PCR *ltmP*
lol 32	TGT CCG TGC ATC CAT TGT	*ltmP*
lol 34	CAT AGA GCT AGC TAG AGT	LHS
lol 35	GTT CGG TGC CTC TAA TAC	*ltmM*/mid
lol 43	GAG GAT AGC CAG CTT GAT	RT-PCR *ltmP*
lol 48	GAT TGG TAC CTT GAA GTC GCT AGT	KO KpnI
lol 49	GTA GGG TAC CTC TAG TAC TGC CTC T	KO KpnI
lol 63	TAG CGA ATC ATT GCG TCG	RT-PCR *ltmP*
lol 79	ATG GCT GCC AAT GAC TTT CC	RT-PCR *ltmG*
lol 135	AGG CCA TTT TCG ACA GTT GT	KO integration
lol 147	CCA GCA AGC ATG CAC ATT AC	RHS
lol 148	TGC GTG AGA GAT AAA GCA AG	KO integration
pUC forward	GCC AGG GTT TTC CCA GTC ACG A	
pUChph 3	CTG CAT CAT CGA AAT TGC	hph
pUChph 4	AAA CCG AAC TGC CCG CTG TTC	hph
PUC reverse	GAG CGG ATA ACA ATT TCA CAC AGG	
T7	TAA TAC GAC TCA CTA TAG GG	

Using degenerate primers designed to fungal GGPP synthase genes, a fragment of the expected size (Fig. 1A) was amplified from lolitrem producing strains, *Neotyphodium lolii* Lp19, and *Epichloë festucae* Fl1, and from the *E. typhina* E8 lolitrem non-producing

strain. *P. paxilli* genomic DNA was used as a positive control where two fragments of 330 and 270 bp were amplified, corresponding to paxG, with an intron, and ggs1, without an intron (Figure 1B). Degenerate PCR amplification was performed using primer pairs ggpps27/ ggpps28 and ggpps27/ ggpps29 with 5 ng of genomic DNA and 4.8 µM of each primer. The amplification conditions were 95°C for 2 min followed by 30 cycles of 95°C for 30 sec, 45°C for 30 sec and 72°C for 1 min, then 1 cycle of 72°C for 5 min. The annealing temperature was also increased to 47°C with a similar amplification result. The resulting products were cloned into pGEM-T easy (Promega). Plasmid DNA was isolated using a BioRad plasmid mini preparation kit. PCR products were purified using a Qiagen PCR purification kit. Fragments were extracted from agarose using the Qiagen gel extraction kit.

The cloned fragments were distinguished using RFLP analysis by amplifying with primers ggpps27 and ggpps28 using standard PCR conditions. The resulting fragments were digested with an appropriate enzyme (*Not*I and *Sau*3AI) and resolved on a 2% agarose gel.

The Lp19 PCR product amplified with primer set ggpps27 and ggpps29 was cloned into pGEM-T easy and sequenced. DNA fragments were sequenced by the dideoxynucleotide chain termination method (Sanger et al. 1977) using Big-Dye (Version 3) chemistry with oligonucleotide primers (Sigma Genosys) to pGEM-T easy, *N. lolii* and *E. festucae* sequences, Products were separated on an ABI Prism 377 sequencer (Perkin-Elmer).

Sequence data was assembled into contigs using SEQUENCHER version 4.1 (Gene Codes) and analyzed using the Wisconsin Package version 9.1 (Genetics Computer Group, Madison, Wisconsin). Sequence comparisons were performed through Internet Explorer version 6.0 at the National Center for Biotechnology Information (NCBI) site

(http://www.ncbi.nlm.nih.gov/) using the Brookhaven (PDB), SWISSPROT and GenBank (CDS translation), PIR and PRF databases employing algorithms for both local (BLASTX and BLASTP) and global (FASTA) alignments (Pearson and Lipman 1988; Altschul et al. 1990; Altschul et al. 1997). A BLASTX of the CY29 sequence, showed high sequence similarity (E value of 7e-41) to the *N. crassa* GGPPS (accession number AAC13867) and other GGPPS sequences. An RFLP screen of the remaining clones revealed a second unique fragment, CY28, that also shows strong similarity to GGPPS genes (the top score was to *P. paxilli Ppggs1* accession number AF279807, Young et al 2001). CY28 was amplified with ggpps27 and ggpps28 and is therefore a shorter product than the CY29 fragment. The two sequences, CY28 and CY29, share 61.7% identity to each other at the DNA level.

To determine which clone is the GGPP synthase involved in lolitrem biosynthesis, each fragment was hybridized to genomic DNA from the two lolitrem producing strains, Lp19 and Fl1, and the non-producing E8 strain. DNA was transferred to positively charged nylon membrane (Roche) using standard techniques (Sambrook et al. 1989). Fragments required for radioactive probes were amplified using primer pairs stated in Table 3. Each probe fragment was purified using a Qiagen PCR purification kit and 30ng of DNA was [α-^{32}P]-dCTP radiolabelled using HighPrime (Roche). The labeled probes were spun through a Pharmacia ProbeQuant column before hybridisation. Hybridisations were performed overnight at 65°C and the filters were washed in 2 x SSC, 0.1% SDS at 50°C.

Table 3 Primer combinations for hybridisation probes and RT-PCR analysis

Gene	primer 1 (5')	primer 2 (3')	Size bp genomic (cDNA)	introns amplified
CY28	g27	g28	209	
CY29	g27	g29	272	
ltmG	lol3	lol1	407 (353)	2
ltmM	lol7	lol35	448 (382)	1
ltmP	lol33	lol37	3277	
ltmP	lol15	lol32	416 (365)	5
ltmG	lol79	lol1	630 (525)	1, 2

ltmM	lol7	lol35	448 (382)	1
ltmM	lol14	lol28	576 (414)	2, 3
ltmP	lol29	lol15	1122 (816)	1, 2, 3, 4, 5
ltmP	lol43	lol63	839 (684)	6, 7

The hybridising patterns (Fig. 1 C & D) showed that CY29 hybridized to all three strains while CY28 hybridised just to the two lolitrem producers, Lp19 and Fl1. This data indicates that CY29 is the orthologue of P. paxilli paxG and CY28 the orthologue of P. paxilli ggs1.

We have named these genes *NlggsA* and *NlltmG* respectively (*ltm* = lolitrem biosynthesis).

Example 2 Isolation of Genomic fragments corresponding to

Using degenerate primers designed to fungal GGPP synthase genes, a fragment of the expected size (Fig. 2A) was amplified from lolitrem producing strains, *Neotyphodium lolii* Lp19, and *Epichloë festucae* Fl1, and from the *E. typhina* E8 lolitrem non-producing strain. *P. paxilli* genomic DNA was used as a positive control where two fragments of 330 and 270 bp were amplified, corresponding to *paxG*, with an intron, and *ggs1*, without an intron (Figure 2B).

The Lp19 PCR product amplified with primer set ggpps27 and ggpps29 was cloned into pGEM-T easy and sequenced. A BlastX analysis of the CY29 sequence, showed high sequence similarity (E value of 7e-41) to the *N. crassa* GGPPS (accession number AAC13867) and other GGPPS sequences (Table 4). An RFLP screen of the remaining clones revealed a second unique fragment, CY28, that also showed strong similarity to GGPPS genes (the top score was to *P. paxilli Ppggs1* accession number AF279807, Young et al. 2001).. CY28 was amplified with ggpps27 and ggpps28 and is therefore a

shorter product than the CY29 fragment. The two sequences, CY28 and CY29, share 61.7% identity to each other at the DNA level.

To determine which clone is the GGPP synthase involved in lolitrem biosynthesis, each fragment was hybridized to genomic DNA from the two lolitrem producing strains, Lp19 and Fl1, and the non-producing E8 strain. The hybridising patterns (Fig. 2 C & D) showed that CY29 hybridized to all three strains while CY28 hybridised just to the two lolitrem producers, Lp19 and Fl1. This data indicates that CY29 is the orthologue of *P. paxilli ggs1* and CY28 the orthologue of *P. paxilli paxG*. We have named these genes *ggsA* and *ltmG* respectively (*ltm* = lolitrem biosynthesis).

The *ltmG* fragment, CY28, was used as a probe to isolate sequences from a Lp19 λGEM12 genomic library. This region of the genome is under represented in the library with only five clones isolated from ~80,000 plated. A 15.6-kb lambda clone, λCY218 (Fig 3), was completely sequenced and shown to contain a complete copy of the *ltmG* gene. To obtain further sequence to the left of *ltmG*, the Lp19 λGEM12 library was screened with a probe amplified with primers lol3 and lol1. Hybridization of the library identified one clone λCY219 that contains extra flanking sequence (Fig. 3), however, this clone was severely rearranged and only 1051 bp reflects the correct genomic arrangement. Sequence analysis of *ltmG* predicts the presence of two introns (Fig 3). These two introns were confirmed by cDNA analysis with RNA isolated from endophyte infected ryegrass. These introns are conserved in position with two of the four introns found in the *ggs-2* gene from *G. fujikouri* (Tudzynski and Hölter 1998) and two of the three introns found in *P. paxilli paxG* (Young et al. 2001). The nucleotide sequence of *ltmG* from *N. lolii* strain Lp19 is shown in Fig. 4. *LtmG* is predicted to encode a polypeptide of 334 amino acids with an unmodified molecular weight of 37.9 kDa (Table 4). The amino acid sequence of the deduced *N. lolii LtmG* polypeptide is shown in Fig. 5. FastA analysis shows that *LtmG* shares 54.1% and 52.6% identity to *N. lolii* GgsA and *P. paxilli* PaxG

polypeptide sequences, respectively. *LtmG* contains the five conserved domains found in all prenyl diphosphate synthases (Chen et al. 1994), including the highly conserved aspartate-rich motifs, DDXXD and DDXXN/D, of domains II and V that are proposed binding sites for the isopentenyl diphosphate (IPP) and the allyl isoprenoid substrates. This analysis suggests that *LtmG* is a GGPP synthase required for the first commited step in lolitrem biosynthesis.

Table 4 Analysis of genes in the lolitrem B biosynthesis cluster

Gene	Putative activity	Size (aa)	Transcript size	Intron number	Homologous *pax* gene	Protein identity
ggsA	Geranylgeranyl diphosphate synthase			0	*ggs1*	
ltmG	Geranylgeranyl diphosphate synthase	334	1002+	2	*paxG*	52.6%
ltmM	FAD dependent monooxygenase	472	1416+	3	*paxM*	41.0%
ltmK	cytochrome P450 monooxygenase	533	1599+	7	*paxP*	31.3%

Example 3. Identification of a gene cluster for lolitrem biosynthesis

Adjacent to *ltmG* are two genes, *ltmM* and *ltmK*, (Fig. 3) proposed to be a FAD-dependent monooxygenase and cytochrome P450 monooxygenase, respectively. Sequence analysis and characterisation by cDNA analysis of the *ltmM* gene shows the presence of three introns (Fig. 3). The first two of these introns are conserved with those found in the P. paxilli *paxM* gene. The third intron is 106 bases, being the largest of the *ltm* introns confirmed. *LtmM* is predicted to encode a polypeptide of 472 amino acids with an unmodified molecular weight of 52.5 kDa (Table 4). The nucleotide sequence of *N. lolii ltmM* and the deduced amino acid sequence of the *LtmM* polypeptide are shown in Figures 6 and 7, respectively. BLASTP analysis showed that *LtmM* shares 41.0% identity

to PaxM from *P. paxilli* (E value 5e-94). Clustal W alignment (Higgins et al. 1994) of *LtmM* with PaxM and other closely related polypeptide sequences, identifies the presence of four highly conserved motifs, the dinucleotide binding domain (Wierenga et al. 1986) the ATG motif (Vallon 2000), a GD motif (Eggink et al. 1990) and a G-helix. These motifs are good indicators of a modified Rossman fold, used by many flavoproteins to bind FAD. This analysis suggests that *LtmM*, like PaxM, is a FAD-dependent monooxygenase, possibly an epoxidase, required for epoxidation of GGPP before cyclisation.

Sequence analysis and characterisation by cDNA analysis of *ltmK* identified seven introns, four of which are conserved with *P. paxilli paxP* and three are conserved with *P. paxilli paxQ*. The nucleotide sequence of *N. lolii ltmK* and the deduced amino acid sequence of the *LtmK* polypeptide are shown in Figures 8 and 9, respectively. *LtmK* is predicted to encode a polypeptide of 533 amino acids with an unmodified molecular weight of 60.9 kDa (Table 4). *LtmK* contains the classical signature motifs of cytochrome P450 enzymes, including a haem-binding domain (Graham-Lorence and Peterson 1996). However, it does not appear to be an orthologue of either PaxP (E value of 9e-62) or PaxQ (E value of 2e-22), two cytochrome P450 enzymes required for paxilline biosynthesis in *P. paxilli* (McMillan et al. 2003), as two other cytochrome P450 genes identified from EST sequences have greater similarity to these genes (see below).

Therefore, *ltmG* forms a gene cluster with an orthologue of *paxM* (*ltmM*) and a cytochrome P450, *ltmK*, of as yet unknown function in lolitrem biosynthesis. The complete nucleotide sequence of this region is shown in Fig. 10. The corresponding region was sequenced from the *E. festucae* strain Fl1 and shown to be 99.9% identical to Lp19, at the DNA level, from the start of *ltmG* to the stop codon of *ltmK*. The nucleotide sequence of *E. festucae ltmG*, *ltmM* and *ltmK* and the deduced amino acid sequence of the corresponding polypeptides LtmG, LtmM and LtmK are shown in

Figures 11 to 16, respectively. Comparison of the *E. festucae ltmM* sequence to *N. lolii ltmM* shows two base transitions of A→G at base ... and T→C at base .... Only the first transition results in a residue change with a conservative replacement of methionine (in *N. lolii ltmM*) to valine (in *E. festucae ltmM*). The promoter region of *N. lolii ltmM* and *E. festucae ltmM* have two differences, the first, T→C at base -356 is at a *Hin*dIII site that is absent from *E. festucae ltmM* and the second is at base -1038 where a GAGA in Lp19 has expanded to GAGAGA in Fl1. *N. lolii ltmK* and *E. festucae ltmK* are identical in sequence.

The DNA sequence flanking the right-hand end of the *ltm* gene cluster contains a high AT content (71.2 %) compared to that of the *ltm* genes at 59.3% AT and, *ggsA* at 40.9% AT. Blast searches of this flanking region reveal sequence similarity to retrotransposons, however, these sequences are very degenerate and no open reading frames are visible.

Example 4 **ltmM *is essential for lolitrem B function* - Deletion of *ltmM* and complementation of *ltmM* mutant**

A gene knockout of *ltmM* in the *E. festucae* strain Fl1 was used to confirm that *ltmM* is essential for lolitrem production. A replacement construct, pCY39, was used in a gene disruption to recombine into the wild-type genome (Fig. 3). An initial PCR screen of 159 hygromycin resistant transformants with primers lol148 and lol135, that amplify both the wild-type *ltmM* gene (1.6 kb) and the integrating plasmid (1.4 kb) identified replacements of *ltmM*. Transformants that contain only the integrating plasmid were 'knockout' candidates and were screened further. The second PCR screen was with primer sets to the upstream (lol2 and lol34: 574 bp), *ltmM* gene (lol7 and lol35: 448 bp), or downstream (lol147 and lol15: 317 bp) regions, where absence of the *ltmM* gene confirmed thea

deletion event. Southern analysis was used to distinguish the true knockouts, of which 3.9% (5/159) contained a single integration of the plasmid. During the screen for a homologous recombination event, a transformant, CYFI1-M61, was identified that has a deletion of *ltmM* and is also deleted beyond *ltmG*, but the extent of the deletion remains uncharacterised.

Two independent knockout strains, CYFI1-M28 (PN2303) and CYFI1-M142 (PN2296), the deletion mutant CYFI1-M61 (PN2301), an ectopic mutant CYFI1-M151 (PN2294), and wild-type FI1 were used to infect endophyte-free perennial ryegrass plants. Each plant was screened for systemic endophyte infection by aniline blue staining confirming normal endophyte associations with the grass. The rate of infection (Table 5) was determined once the plants had established reasonable growth and shows that each strain has a similar infection rate. The endophyte infected plants were grown in a containment green house and were screened for alkaloid production in mid-summer. The alkaloid levels (summarised in Table 5) show that the endophyte strains with a deleted *ltmM* gene (CYFI1-M28, CYFI1-M61 and CYFI1-M142) are unable to produce lolitrem B (Fig. 17) but the level of ergovaline and peramine production is consistent with wild-type and the ectopic integrant CYFI1-M151. The knockouts (CYFI1-M28 and CYFI1-M142) and deletion mutant (CYFI1-M61) are also devoid of two smaller peaks, that are assumed to be lolitrem A and E, respectively (Fig. 17).

A complementation construct for *ltmM*, pCYItmM, was made by cloning a 7 -kb *Xho*I fragment containing 2.2 kb of 5' and 3kb of 3' *ltmM* sequences into pII99. Four random integrants of PN2303 containing this construct were infected into plants and shown to synthesize lolitrems.

Plant Inoculation

Two independent knockout strains, CYFI1-M28 and CYFI1-M142, the deletion mutant CYFI1-M61, an ectopic mutant CYFI1-M151, and wild-type FI1 were used to artificially infect endophyte-free perennial ryegrass plants. Ryegrass cultivar Nui was infected with fungal endophyte according to the procedure of (Latch and Christensen 1985). Four - five weeks after inoculation the plants were checked for systemic endophyte infection by immunoblotting with endophyte antisera and staining pseudostem leaf peels with aniline blue to detect the presence of the endophyte. Plants that were endophyte positive were repotted and allowed to grow under greenhouse conditions. The rate of infection (Table 5) was determined once the plants had established reasonable growth and shows that each strain has a similar infection rate.

Alkaloid Analysis

The endophyte infected plants were grown in a containment green house and were screened for alkaloid production mid-summer. Endophyte infected plant pseudostem material was freeze dried and milled. For lolitrem analysis weighed portions (c. 50 mg) were extracted for 1 hour at ambient temperature with 1 ml of dichloroethane-methanol, 9:1 by volume, in 2 ml polypropylene screw cap vials turning end for end for agitation. The extract was separated by centrifugation and 8 µl portions were examined for lolitrems by normal phase high performance liquid chromatography (Shimadzu LC-10A system) on Alltima silica 5µ 150 x 4.6 mm columns (Alltech Associates, Deerfield, Il). The mobile solvent was dichloromethane-acetonitrile-water, 860:140:1 by volume, with a flow rate of 1 ml/min. Lolitrems were detected by fluorescence (Shimadzu RF-10A, excitation 265 nm, emission 440 nm). Lolitrem B eluted at approximately 4.5 minutes followed by smaller amounts of other lolitrems. The amount of lolitrem B was estimated by comparison of integrated peak areas with external standards of authentic lolitrem B. The detection limit was estimated as < 0.1 ppm of lolitrem B.

Ergovaline and peramine were analysed by the method of Spiering et al.(2002). (reference is Spiering MJ, Davies E, Tapper BA, Schmid J, Lane GA (2002) Simplified extraction of ergovaline and peramine for analysis of tissue distribution in endophyte-infected grass tillers. J Agricultural and Food Chem 50:5856-5862.

The alkaloid levels (summarised in Table 5) show that the endophyte strains with a deleted *ltmM* gene (CYFI1-M28, CYFI1-M61 and CYFI1-M142) are unable to produce lolitrem B (Fig. 5) but the level of ergovaline and peramine production is consistent with wild-type and the ectopic integrant CYFI1-M151. The knockouts (CYFI1-M28 and CYFI1-M142) and deletion mutant (CYFI1-M61) are also devoid of two smaller peaks, that are assumed to be lolitrem A and E, respectively (Fig. 5).

Table 5 Rates of infection, fungal biomass and alkaloid production

Strain	Fungal Type [1]	Number of plants/ association	Infection Rate[2] (%)	Lolitrem (ppm)	Ergovaline (ppm)	Peramine (ppm)
CYFI1M-28	KO	5	20	0	0.4 - 1.3	30 - 40
CYFI1M-61	Del	4	17	0	0.7 - 3.3	24 - 41
CYFI1M-142	KO	5	17	0	0.1 - 2.0	14 - 47
CYFI1M-151	Ectopic	5	17	4.4 - 16.7	0.5 - 1.2	21 - 55
FI1	Wt	4	22	6.2 - 12.8	0.8 - 1.5	31 - 66
Endophyte Free	NA	3	NA	0	0	0

[1]KO = *ltmM* knockout, Del = deletion mutant, Wt = Wildtype, NA = Not applicable.

[2]Infection rates were determined as a percentage of endophyte infected from the surviving plants. The infection rates are low as the endophyte is inserted into young plants at a wound site.

Example 5 Construction and sequencing of Suppressive Subtractive Hybridisation Libraries

To identify additional genes involved in the lolitrem biosynthetic pathway an approach of EST sequencing from both *N. lolii in vitro* culture derived cDNA libraries and from subtracted plant derived cDNA libraries was adopted. ESTs within the libraries derived from *N. lolii* and with homology to genes from the paxilline biosynthetic pathway are good candidates for orthologous lolitrem biosynthetic genes. It was expected that some genes may be expressed in *in vitro* cultures but many may only be expressed *in planta* so the dual approach was taken. The *in vitro* culture libraries were derived from liquid cultures in both rich and minimal media to increase the chance of identifying ESTs that may only be expressed under starvation conditions and are described in example 6. The subtracted libraries were derived by constructing cDNA from both infected and uninfected perennial ryegrass plants and performing suppressive subtractive hybridization to enrich for fungal cDNAs.

Infected Plant Material

Perennial ryegrass genotypes are genetically complex due to the outbreeding nature of this species. To eliminate plant genotype effects and enable the comparison of infected and uninfected perennial ryegrass plants with identical genetic backgrounds cloned lines of infected Nui were cured of the fungus. The isogenic ryegrass lines infected or uninfected with N. lolii strain Lp19 (AR42) were produced by Mike Christnesen (AgResearch Limited). Lp19 (AR42) is a novel endophyte from the AgResearch

collection. Produces Lolitrem B, Ergovaline and Peramine. Lp 19 is an endophyte that has been isolated from its parent plant and inoculated into the ryegrass cultivar Nui i.e. it is a novel endophyte.

Positive and negative clones of the above material were produced by taking a positive plant and dividing the tillers up to produce a number of cloned plants. Some of the clones were then treated with a systemic fungicide to eliminate the endophyte.

This was done by striping tillers down and soaking in a 2g/L solution of Benlate (50% Benomyl w/w) for several hours then planting them in clean river sand saturated with the solution. Pots were watered to weight for several weeks such that the tillers were essentially immersed in fungicide for this period. Plants are potted into commercial potting mix and tillers assayed for endophyte presence. Endophyte free tillers were removed to new pots and tested periodically for endophyte presence to ensure that the fungus has been successfully eliminated. In this way we obtain E+ and E- cloned copies of an individual ryegrass genotype.

Plants were grown in the glasshouse in pots containing commercial potting mix. Plants were dissected in order to provide emerging immature leaf tissue and mature sheath tissue. Material was harvested and frozen immediately at -80C until needed.

Development of Suppressive Subtractive Hybridisation Libraries

RNA was extracted from the harvested plant tissues using the Triazol method (Invitrogen) following the manufacturers recommendations. Messenger RNA was purified from this using mRNA purification kits (Amersham) following the manufacturers recommendations. Messenger RNA (mRNA) was used in subsequent subtractive hybridisations using the Suppressive Subtractive Hybridisation (SSH) kit (Clontech) as per the manufacturer's instructions.

Subtractions were carried out in both a 'forward' and 'reverse' direction using 'tester' and 'driver' cDNAs as follows:

Tester equals cDNA from infected plants (Nle+).

Driver equals cDNA from uninfected plants (Nle-).

Plant line	Leaf tissue	Library
Nle+M	Mature	Up-regulated
Nle-M	Mature	Down-regulated
Nle+I	Immature	Up-regulated
Nle-I	Immature	Down-regulated

Subtractions were carried out using tester and driver from both immature and mature tissue and in both directions. Forward subtractions enrich for up-regulated genes and reverse subtractions enrich for down-regulated genes. After the subtraction procedure, cDNAs were ligated into the vector pCR-Topo2.1 (Invitrogen) and transformed into *E. coli* competent cells following the manufacturers recommendations. 1000 clones from each library were stored as glycerols in 96 well format.

Template preparation and Library sequencing

For sequencing template preparation PCR reactions were carried out in 384-well plates using the M13 forward (GTAAAACGACGGCCAG) and Reverse primers (CAGGAAACAGCTATGAC). The Biomek 2000 liquid handling robot was used to transfer 1 µl aliquots from each of 4 x 96-well plates containing overnight cultures into a

conical bottomed 384-well plate (ABGen). PCR products were precipitated using 1 µl of 3M NaoAC (pH 6) and 15 µl of isopropanol and placed at -80°C for at least one hour before centrifugation at 4K for 1 hr (4°C). Pellets were washed with 20 µl of 70% ethanol and centrifuged for a further 30 min at 4K (4°C) before they were air dried and resuspended in 10 µl of sterile MQ water. Products were checked by running 1 µl samples on a 1% agarose gel (1X TAE).

Sequencing reactions were performed in conical bottomed 384-well plates (Applied Biosystems) using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). One µl of PCR product was added to 9 µl of sequencing mix (0.8 µl of 2 µM M13 Reverse primer; 0.5 µl Big Dye (Version 3); 3.5 µl ABI dilution buffer (400 mM Tris pH9; 10 mM $MgCl_2$) and 4.2 µl sterile MQ water) and the plate centrifuged briefly to collect the contents at the bottom of the wells. Cycle sequencing was performed using 40 cycles of 95°c for 20 sec, 50°C for 15 sec and 60°c for 1.5 min (iCycler, Bio-Rad, USA). Sequencing products were precipitated by the addition of 1 µl of 3M NaOAc (pH 4.6), 1 µl sterile MQ water and 23 µl non-denatured 95% ethanol, placed on the bench for 15 min (at RT) and centrifuged at 4K for 30 min (4°C). Immediately following centrifugation, plates were turned upside down on to several sheets of paper towels and centrifuged at 50 x *g* for 1 min to expel all liquid. Any remaining liquid was removed by briefly spinning the plate in a salad spinner and the pellets resuspended in 10 µl of HiDi™ formamide solution (Applied Biosystems). Sequencing was performed on the ABI 3100 (Applied Biosystems) using a 36 cm array.

Example 6. Construction of EST Database from *in vitro* Cultures

To identify additional genes involved in the lolitrem biosynthetic pathway an approach of EST sequencing from *N. lolii in vitro* culture derived cDNA libraries was adopted. ESTs

within the libraries derived from *N. lolii* and with homology to genes from the paxilline biosynthetic pathway are good candidates for orthologous lolitrem biosynthetic genes. It was expected that some genes may be expressed in *in vitro* cultures but many may only be expressed *in planta* so an *in planta* approach is described in example 5. The *in vitro* culture libraries were derived from liquid cultures in both rich and minimal media to increase the chance of identifying ESTs that may only be expressed under starvation conditions.

Culture Conditions

N. lolli strain Lp19 was initially cultured on potato dextrose agar plates. Mycelia from the leading edge of colonies were removed and chopped up finely with a scalpel blade before being transferred to 50 ml potato dextrose broth and incubated for 10 days at 25°C/200 rpm. Mycelia for RNA extraction were harvested under vacuum using a sterile buchner funnel containing two layers of Whatman 3 MM paper. Mycelia were washed three times with sterile MQ water. Dry weight was estimated from the wet weight.

To grow mycelia in minimal media, mycelia from *N. lolli* strain Lp19 cultures initially grown in complete medium for 14 days were harvested under vacuum using a sterile buchner funnel containing two layers of Whatman 3 MM paper. Mycelia were washed three times with sterile MQ water before transfer to the minimal medium. Two grams of mycelia was used to inoculate 50 ml of Blankenship MM and the cultures incubated for 19 days at 25°C/200 rpm. Mycelia for RNA extraction were harvested under vacuum using a sterile buchner funnel containing two layers of Whatman 3 MM paper. Mycelia were washed three times with sterile MQ water. Dry weight was estimated from the wet weight.

Isolation of total RNA from cultures grown in complete medium

Mycelia were ground in liquid nitrogen using a sterile mortar and pestle and the RNA extracted using 1 ml Trizol® Reagent (Invitrogen) per 100 mg of ground mycelia. Samples were mixed well and frozen overnight at -20°C. The following day samples were thawed at RT on an orbital mixer (approx. 1 hr) and centrifuged at 12000 *x* g for 10 min (4°C) to remove polysaccharides. The supernatant was removed to a fresh tube and 0.2 ml of chloroform added per 1 ml Trizol reagent. Tubes were capped well and shaken vigorously by hand for 15 s and incubated at RT for 2 to 3min. Samples were centrifuged at 12000 *x* g for 15 min at 4°C and the supernatant removed with a pipette to a fresh tube. RNA was precipitated using a modified precipitation step that effectively precipitated the RNA while maintaining polysaccharides and proteoglycans in a soluble form. Essentially, 0.25 ml isopropanol was added to the supernatant followed by 0.25 ml of a high salt precipitation solution (0.8M sodium citrate and 1.2M NaCl) per 1 ml of Trizol reagent used for the initial homogenization. The resulting solution was mixed well and the samples incubated at RT for 10 min. Samples were centrifuged at 12 000 *x* g for 10 min at 4°C and the resulting RNA pellet washed once with 75% ethanol (1 ml 75% ethanol per 1 ml Trizol). The sample was mixed by vortexing and centrifuged at 7500 *x* g for 5 min at 4°C.

The RNA pellet was briefly air dried for 5-10 min at RT and dissolved in 1 ml RNase free water (Invitrogen) with 1 µl Protector RNase (Roche) by passing the solution several times through a pipette tip and incubating for 10 min at 55-60°C. RNA purity and concentration were determined by spectrophotometry ($A_{260/280}$) and by running 3 µl and 6 µl aliquots (containing 1 µl of 10X MOPS running dye (0.2 M MOPS (pH7), 20 mM sodium acetate, 10 mM EDTA (pH8) in a total volume of 10 µl) on a 1% agarose gel in 1 X TAE buffer containing ethidium bromide (1 µg/ml). RNA was stored as 10 µl aliquots at -80°C.

Isolation of total RNA from cultures grown in minimal medium

Mycelia were ground in liquid nitrogen using a sterile mortar and pestle and the RNA extracted using 1 ml Trizol® Reagent (Invitrogen) per 100 mg of ground mycelia. Samples were mixed well and frozen overnight at -20°C. The following day samples were thawed at RT on an orbital mixer (approx. 1 hr) and 0.2 ml chloroform added per 1 ml of Trizol reagent. Samples were vigorously shaken by hand for 15 s and incubated at RT for 2-3 min. Samples were centrifuged at 12 000 *x* g for 15 min at 4°C and the upper aqueous phase removed to a fresh tube. RNA was precipitated using 0.5 ml isopropanol per 1 ml Trizol reagent used for the initial homogenization. Samples were incubated at RT for 10 min and centrifuged at 12 000 *x* g for 10 min at 4°C. The RNA pellet was washed using 1 ml 75% ethanol per 1 ml Trizol reagent used for the initial homogenization. Samples were mixed by vortexing and centrifuged at 7500 *x* g for 5 min at 4°C. The RNA pellet was briefly air dried for 5-10 min at RT and dissolved in 1 ml RNase free water (Invitrogen) with 1 µl Protector RNase (Roche) by passing the solution several times through a pipette tip and incubating for 10 min at 55-60°C. RNA purity and concentration were determined by spectrophotometry ($A_{260/280}$) and by running 3 µl and 6 µl samples on a 1% agarose gel in 1 X TAE buffer containing ethidium bromide. RNA was stored as 10 µl aliquots at -80°C.

Purification of mRNA

mRNA was purified from total RNA using the mRNA Purification Kit (Amersham Pharmacia Biotech) as per the manufacturer's instructions. Each Oligo (dT)-cellulose column had the capacity to bind approximately 25 µg of poly(A)$^+$RNA so, assuming that only 2% of the total RNA was polyadenylated, no more than 1.25 mg of total RNA was applied to each column. mRNA was subjected to two rounds of purification and the concentration determined by spectrophotometry ($A_{260/280}$). Aliquots were stored at -80°C.

First-strand cDNA synthesis using mRNA

Two µl of mRNA was combined with 1 µl SMART IV™ oligonucleotide and 1 µl CDS III/3' PCR primer in a sterile thin-walled 0.2 ml PCR tube (Bio-Rad). The contents were mixed, spun briefly in a microfuge and incubated at 72°C for 2 min. The tube was then cooled on ice for 2 min, spun briefly to collect the contents at the bottom of the tube and the following added:

2 µl 5X First-Strand Buffer

1 µl DTT (20 mM)

1 µl dNTP mix (10 mM)

1 µl PowerScript™ Reverse Transcriptase

Samples were mixed by gentle pipetting, briefly spun to collect the contents and incubated at 42°C for 1 hr (Bio-Rad iCycler). The tube was placed on ice to terminate first-strand synthesis, 1 µl of sodium hydroxide (25 mM) added and the tube incubated at 68°C for 30 min. A 3 µl aliquot was removed for cDNA amplification by Primer Extension PCR and the remaining first-strand cDNA stored at -20°C.

First-strand cDNA synthesis using total RNA

Three µl of freshly-prepared total RNA was combined with 1 µl SMART IV oligonucleotide and 1 µl CDS III/3' PCR primer in a sterile thin-walled 0.2 ml PCR tube (Bio-Rad). The contents were mixed, spun briefly in a microfuge and incubated at 72°C for 2 min. The tube was then cooled on ice for 2 min, spun briefly to collect the contents at the bottom of the tube and the following added:

2 µl 5X First-Strand Buffer

1 µl DTT (20 mM)

1 µl dNTP mix (10 mM)

1 µl PowerScript™ Reverse Transcriptase

Samples were mixed by gentle pipetting, briefly spun to collect the contents and incubated at 42°C for 1 hr (Bio-Rad iCycler). The tube was placed on ice to terminate first-strand synthesis and a 3 µl aliquot removed for cDNA amplification by Long Distance (LD) PCR. The remaining first-strand cDNA was stored at -20°C.

cDNA amplification by Primer Extension PCR

The following components were combined in a sterile 0.2 ml thin-walled PCR tube:

11 µl	First Strand cDNA (from section 3.1)
71 µl	sterile MQ water
10 µl	10X Advantage 2 PCR buffer
2 µl	50X dNTP mix
2 µl	5' PCR primer
2 µl	CDS III/3' PCR primer
2 µl	10X Advantage 2 Polymerase mix

Samples were mixed, briefly spun to collect the contents and amplified by PCR (72°C for 10 min, 95°C for 20 s and 3cycles of 95°C for 5 s, 68°C for 8 min) using the Bio-Rad iCycler. A 10 µl sample was analysed on a 1.0% agarose gel (1X TAE) alongside 0.1 µg of a 1 kb plus DNA size marker (Invitrogen). The ds cDNA either underwent subsequent proteinase K and *Sfi* I digestions or was stored at -20°C until further use.

cDNA amplification by LD PCR

The following components were combined in a sterile 0.2 ml thin-walled PCR tube (Bio-Rad):

3 µl	First-Strand cDNA (from section 3.2)
79 µl	sterile MQ water
10 µl	Advantage 2 PCR buffer
2 µl	50X dNTP mix
2 µl	5' PCR Primer
2 µl	CDS III/3' PCR Primer
2 µl	50X Advantage 2 Polymerase Mix

Samples were mixed by gently flicking the tube, briefly spun to collect the contents and amplified by PCR (95°C for 30 s and 26 cycles of 95°C for 15 s, 68°C for 6 min) using the Bio-Rad iCycler. The ds cDNA either underwent subsequent proteinase K and *Sfi* I digestions or was stored at -20°C until further use.

Four µl of Proteinase K (20 µg/µl) and 5 µl of sterile MQ water were added to 90 µl of amplified ds cDNA, mixed and incubated at 45°C for 20 min. The reaction was cleaned up using the Qiagen PCR Purification Kit as per the manufacturer's instructions and the cDNA eluted from the column in a total volume of 50 µl.

The following components were added to a fresh 0.2 ml thin-walled PCR tube:

50 µl	cDNA (proteinase K treated)

29 µl sterile MQ water

10 µl 10X *Sfi* I buffer

10 µl *Sfi* I restriction enzyme

1 µl 100X BSA

Samples were mixed well and incubated at 50°C for 2 hr.

Following *Sfi* I digestion, 2 µl of a 1% xylene cyanol solution was added to the tube and the sample mixed well. Sixteen sterile 1.5 ml tubes were labelled and arranged in a rack in order. A CHROMA SPIN-400 column (Clontech) was prepared as per the manufacturer's instructions and the mixture of Sfi I-digested cDNA and xylene cyanol dye carefully applied to the top centre surface of the column matrix. Once the sample was fully absorbed into the matrix, 100 µl of column buffer was also applied to the column and the buffer allowed to drain from the column until there was no liquid remaining above the resin. At this point, the dye layer was several mm into the column.

The rack containing the 1.5 ml collection tubes was placed so that the first tube was directly underneath the column outlet. 600 µl of column buffer was added to the column and single-drop fractions (approximately 35 µl per tube) collected in the labelled tubes. The profile of each fraction was checked by analysing 10 µl samples alongside 0.1 µg of a 1 kb plus DNA standard (Invitrogen) on a 1.1% agarose gel (1X TAE; 150V; 10 min). The gel was stained with ethidium bromide for 15 min, destained in water for 1.5 hr and the peak fractions determined by visualizing the intensity of the bands under UV. The first 3 fractions containing cDNA were collected and pooled. Samples were cleaned up using an Amicon-30 unit (Millipore). The unit was washed twice with sterile MQ water before use as per the manufacturer's instructions. The pooled fractions were applied to the unit and concentrated to 7 µl by centrifugation at 14 000g for 20 min at room

temperature. The *Sfi* I-digested cDNA was either stored at -20°C or used immediately in the ligation reaction.

Ligation of Sfi I-digested cDNA to the ?TriplEx2 Vector and library packaging

Ligations were optimized using three different ratios of cDNA to phage vector following the manufacturers recommendations. Samples were mixed gently, centrifuged briefly to bring the contents to the bottom of the tube and incubated overnight at 16°C. Ligations (cDNA/?TriplEx2 Vector) were heat inactivated at 65°C for 15 min. Packaging reactions (50 µl) were thawed at room temperature and placed on ice. Half of the packaging extract (25 µl) was immediately transferred to a second ice-cold 1.5 ml tube. The entire ligation (7 µl) was added to 25 µl of packaging extract, mixed gently with a pipette and incubated at 30°C for 90 min. At the end of this incubation, the remaining 25 µl of packaging extract was added to the sample and the reaction incubated for a further 90 min at 30°C. Five hundred µl of 1X Lambda dilution buffer (100 mM NaCl, 10 mM $MgSO_4.7H_2O$, 35 mM Tris-HCl (pH 7.5), 0.01% gelatin) was added to the sample and mixed by gentle vortexing. Chloroform (25 µl) was also added to prevent bacterial contamination. Packaged libraries were titered following the manufacturers recommendations and stored at 4°C for up to one month.

Library Amplification

A single, well-isolated colony of XL1-Blue was picked from the primary working plate and used to inoculate 15 ml of LB broth containing $MgSO_4$ (10 mM) and maltose (0.2%). Cultures were incubated at 37°C overnight with shaking (140 rpm). Cells were harvested the following day by centrifuging the culture at 5K for 5 min. The supernatant was removed by decanting and the pellet resuspended in 7.5 ml of 10 mM $MgSO_4$. Enough phage to yield 6-7 x 10^4 plaques per 150 mm plate was added to each of 10 tubes containing 500 µl of overnight XL1-Blue culture in a sterile 1.5 ml tube. Phage

were allowed to adsorb to the *E. coli* cells by incubating in a 37°C water bathfor 15 min before adding 4.5 ml of melted (45°C) LB top agar containing $MgSO_4$ (10 mM) and maltose (0.2%). Samples were quickly mixed by gentle vortexing and immediately poured on to prewarmed (37°C) 150 mm LB agar plates containing $MgSO_4$ (10 mM). Plates were cooled for 10 min at room temperature to allow the top agar to harden and incubated at 37 °C for 10.5 hr. Phage were eluted by adding 12.5 ml of 1X Lambda dilution buffer to each plate and the plates stored overnight at 4°C. The following day, the plates were shaken (~50 rpm) at room temperature for 1 hr and the phage lysates poured into a sterile beaker. Intact cells were lysed by adding 10 ml of chloroform and the phage lysate cleared of cell debris by centrifuging at 5 000 *x* g for 10 min in sterile 50 ml polypropylene tubes. The supernatant was collected and stored at 4°C in sterile universals. For long-term storage, 1 ml aliquots were made containing DMSO to a final concentration of 7% and frozen at -80°C.

Converting ?TriplEx2 to pTriplEx

The bacterial host strain *E. coli* BM25.8 (*supE*44, thi ?(lac-proAB) *relA*1, [F' *lacI*^{q}Z?*M15*, *proAB*$^+$, *traD*36], *hsdR*(r_{k12}-m_{k12}-), (*kan*^{R})P1 (*cam*^{R}) ?*imm*434) was supplied as a component of the SMART cDNA Library Construction Kit (Clontech) and stored at -80°C. For large-scale library conversion a single, well-isolated colony of *E. coli* BM25.8 was picked from the primary working plate and used to inoculate 10 ml of LB broth. Cultures were incubated at 31°C overnight with shaking (150 rpm). The following day, $MgCl_2$ (10 mM) was added to the overnight culture of BM25.8. In a sterile 15 ml tube, 200 µl of overnight culture was mixed with 2 x 10^6 pfu/ml of amplified ?TriplEx2 cDNA library and incubated for 1 hr at 31°C (without shaking). After the incubation was complete, 500 µl of LB broth was added and the sample incubated for a further 1 hr at 31°C with shaking (190 rpm). At this point, conversion of the library to plasmid form was complete. The converted cDNA library was diluted 1:100 in LB broth and aliquots (10 µl, 100 µl) were

spread on to LB agar plates containing carbenicillin (50 µg/ml). Plates were incubated overnight at 31°C and the colonies picked for further analysis. The remaining converted library was stored as 1 ml aliquots containing glycerol (to a final concentration of 30%) at -80°C.

PCR analysis

Individual colonies from converted libraries were inoculated into 100 µl of LB broth containing carbenicillin (50 µg/ml) in round bottomed 96-well plates (Nunc). Plates were incubated overnight at 37°C. Aliquots of 1 µl of each overnight culture were PCR amplified in a total volume of 15 µl using ptriplex2FORWARD (5'-AAGCGCGCCATTGTGTTGGTACCC-3') and ptriplex2REVERSE (5'-CGGCCGCATGCATAAGCTTGCTCG-3') as primers (present in the pTriplEx vector arms) (Kohler *et al.*, 2003). The PCR included 95°C for 3 min, 95°C for 60 s, 60°C for 30 s, 72°C for 3 min for 30 cycles and a final extension of 72°C for 15 min (iCycler, Bio-Rad, USA). One µl of each reaction was analysed on a 1% agarose gel alongside 0.25 µg of a 1 kb plus DNA standard (Invitrogen) and stained with ethidium bromide to determine the size and quality of the PCR products.

For sequencing template preparation PCR reactions were carried out in 384-well plates. The Biomek 2000 liquid handling robot was used to transfer 1 µl aliquots from each of 4 x 96-well plates containing overnight cultures into a conical bottomed 384-well plate (ABGen). PCR products were precipitated using 1 µl of 3M NaoAC (pH 6) and 15 µl of isopropanol and placed at -80°C for at least one hour before centrifugation at 4K for 1 hr (4°C). Pellets were washed with 20 µl of 70% ethanol and centrifuged for a further 30 min at 4K (4°C) before they were air dried and resuspended in 10 µl of sterile MQ water. Products were checked by running 1 µl samples on a 1% agarose gel (1X TAE) and

further diluted either 1:5 (minimal medium cDNA library) or 1:1 (complete medium cDNA library) in sterile MQ water before sequencing.

Sequencing Reactions

Sequencing reactions were performed in conical bottomed 384-well plates (Applied Biosystems) using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems). One µl of PCR product was added to 9 µl of sequencing mix (0.8 µl of 2 µM ptriplex2FORWARD primer; 0.5 µl Big Dye (Version 3); 3.5 µl ABI dilution buffer (400 mM Tris pH9; 10 mM $MgCl_2$) and 4.2 µl sterile MQ water) and the plate centrifuged briefly to collect the contents at the bottom of the wells. Cycle sequencing was performed using 40 cycles of 95°c for 20 sec, 50°C for 15 sec and 60°c for 1.5 min (iCycler, Bio-Rad, USA). Sequencing products were precipitated by the addition of 1 µl of 3M NaOAc (pH 4.6), 1 µl sterile MQ water and 23 µl non-denatured 95% ethanol, placed on the bench for 15 min (at RT) and centrifuged at 4K for 30 min (4°C). Immediately following centrifugation, plates were turned upside down on to several sheets of paper towels and centrifuged at 50 x *g* for 1 min to expel all liquid. Any remaining liquid was removed by briefly spinning the plate in a salad spinner and the pellets resuspended in 10 µl of HiDi™ formamide solution (Applied Biosystems). Sequencing was performed on the ABI 3730 (Applied Biosystems) using a 50 cm array.

Example 7. Identification of ESTs Encoding Putative Lolitrem Biosynthetic Genes from EST Sequence Databases

A sequence database was developed containing the 4000 EST sequences derived from the SSH libraries and 6500 ESTs derived from *in vitro* culture libraries. The database was searched using the BLAST algorithm. Nucleotice sequences were blasted using the

BlastX algorithm against the SwissProt database. ESTs with homology to paxilline biosynthetic genes are listed in table 6. All paxilline orthologs were identified in the Nle+M library.

Table 6. Detail of ESTs with Homology to Paxilline Biosynthetic Genes

EST	Length (bp)	Function	Paxilline Homolog	Blast Score
E07	353	dimethylallyltryptophan (DMAT) synthase	paxD	5e-02
DMAT Johnson1	532	dimethylallyltryptophan (DMAT) synthase	paxD	
N17	413	Cytochrome P450 monooxygenase		2e-09
G13	335	Cytochrome P450 monooxygenase	paxP	3e-07
J15	639	Cytochrome P450 monooxygenase	paxP	8e-34

Example 8. Predicted Genes in Lolitrem Gene Cluster

Isolation of Lolitrem Biosynthetic Genes

We describe here the molecular cloning and genetic analysis of a set of genes from *N. lolii* and *E. festucae* that are proposed to be involved in the biosynthesis of lolitrem and closely related indole-diterpenes. This is the second indole-diterpene gene cluster to be cloned from a filamentous fungus. We recently reported on the isolation of a cluster of genes from *P. paxilli* required for paxilline biosynthesis (Young et al. 2001). A comparison with the paxilline biosynthesis cluster identifies five functional orthologues, *ltmG*, *ltmM*, *ltmP*, *ltmQ* and *ltmD*. In addition we have identified two P450 genes, *ltmJ* and *ltmK*, that may also play a role in indole-diterpene biosynthesis in this group of fungi. Three of these genes, *ltmG*, *ltmM* and *ltmK* form a tightly linked cluster.

The first of these genes, *ltmG*, is clearly identifiable as a GGPP synthase, and is presumed to catalyse the first step in the biosynthesis of lolitrems i.e. the synthesis of GGPP. Interestingly, the two fungal species in which diterpene gene clusters have been analysed, have two copies of GGPP synthase, one proposed to be involved in primary metabolism and one specifically recruited for secondary metabolism (Tudzynski and Hölter 1998; Young et al. 2001). *N. lolii* and *E. festucae* also have two copies of a GGPP synthase. One copy is proposed to be required for primary metabolism and the second copy (*ltmG*) is proposed to be specifically required for indole-diterpene biosynthesis.

Deletions of *paxM* and *paxC* in *P. paxilli* result in mutants with a paxilline-negative phenotype. To date no identifiable indole-diterpene intermediates have been identified in these strains, suggesting that these genes are involved in very early steps in the pathway. Our working model is that PaxM and PaxC are required to catalyse the epoxidation and cyclisation of GGPP and addition of indole-3-glycerol to form the first stable indole-diterpene, possibly paspaline (Parker and Scott 2004). By analogy we propose that LtmM catalyses the same early reaction in lolitrem biosynthesis. In support of this hypothesis we were able to demonstrate that *ltmM* is required for lolitrem biosynthesis by making a targeted deletion of this gene. Mutants deleted in this gene were unable to synthesize lolitrem B in artificial symbiota with perennial ryegrass. An *N. lolii* orthologue of *paxC*, is yet to be identified, but is predicted to also be essential for lolitrem biosynthesis.

Other genes identified as being necessary for paxilline biosynthesis are *paxP* and *paxQ*; which encode cytochrome P450 enzymes. Targeted deletion of *paxP* and *paxQ* results in strains that accumulate paspaline and 13-desoxypaxilline, respectively (McMillan et al. 2003). These results suggest that PaxP is required for demethylation of C-12 of paspaline, and possibly hydroxylation of C-10, and PaxQ is required for hydroxylation of C-13, using either PC-M6 or 13-desoxypaxilline as substrates (Fig. 16). Analysis of the

structure of lolitrem B (Fig. 1) suggests that similar modifications are required to the paspaline skeleton (Fig. 16) to generate lolitrem B. Orthologues of *paxP* and *paxQ* were identified in an EST library generated with template from suppression subtractive hybridization. A further enzyme predicted to be required for lolitrem B biosynthesis is a prenyl transferase to prenylate positions 20 and 21 of the indole ring. A candidate gene for one or both of these prenylations is *ltmD*, given that the paralogue, *dmaW*, prenylates position 20, as the first committed step in ergot alkaloid biosynthesis (Wang et al. 2004). One or two additional cytochrome P450 enzymes are predicted to be required for further oxidation and closure of ring A of lolitrem B. Candidates for these functions include *ltmJ* and *ltmK*. At least two additional enzymes are required to form an epoxide between C-11 and C-12 of paspaline, and prenylate ring H of lolitrem B. These would be predicted to be an FAD-dependent monooxygenase and a prenyl transferase, respectively. We have yet to identify these genes.

In summary, we predict that up to ten genes are required for the biosynthesis of lolitrem B. Candidate genes identified to date include *ltmG*, *ltmM*, *ltmK*, *ltmP*, *ltmQ*, *ltmD* and *ltmJ*. Deletion analysis has confirmed that at least *ltmM* is required for lolitrem B biosynthesis. Further genetic analysis of the genes identified here and adjacent genes will help elucidate the pathway for lolitrem biosynthesis. A comparison with the steps required for paxilline biosynthesis in *P. paxilli* will elucidate the basic biochemistry and genetics of this important group of secondary metabolites.

Example 9. Demonstration that the ESTs with homology to *Penicillium paxili* paxP do infact belong to the lolitrem biosynthetic gene cluster can be done by a number of approaches. For example the EST can be used as a probe to screen the genomic library described in example 2. Clone with homology to the probe can be isolated and sequenced. Bioinformatic analysis of the sequence will determine if any other gene candidates are present on the lambda clones. Confirmation that these genes are linked

to ltmM, ltmG and ltmK can be made be probing a CHEF gel containing *N. lolii* high molecular weight DNA. If the isolated genes hybridise the the same high molecular weight DNA band as ltmG then the genes are linked and probably part of the same gene cluster. Final conformation or the role of these candidate genes can be made by generating a mutant endophyte strain by homologous recombination where the gene is deleted. The mutant is expected to be lolitrem minus.

Example 10. Methods for Expression of Lolitrem genes in Transgenic Plants

Once the lolitrem biosynthetic gene cluster is fully characterized it is possible to modify the fungal genes to enable expression in transgenic plants. Fungal genes containing introns will not be correctly spliced in plants so cDNAs for each gene need to be obtained. Those familiar with the art will know it is possible to isolate cDNAs using cDNA synthesis kits such as those described in example 6. The cDNAs need to be cloned into a vector that contains a plant promoter and terminator sequence. Those familiar with the art know that there are many possible promoter and terminator combinations. A common example is the 35s promoter from Cauliflower Mosaic Virus (Odell et al., 1985). These modified fungal genes can then be transformed into plant species using either the gene gun or agrobacterium. Two methods are described below.

Transformation of Lolium perenne by Microprojectile bombardment of embryogenic callus

It is possible to use perennial ryegrass L. perenne as a model system for monocot plant species. Demonstration of biosynthesis of indole diterpenes in this species can be extrapolated to other monocot species such as wheat, rice and corn.

Materials

- florally induced tillers of *Lolium perenne*
- Na-hypochlorite (5% available chlorine)
- sterile ddH_2O100mm Petri plates containing LP5 medium*
- 100mm Petri plates containing LP3-OS medium
- 100mm Petri plates containing LP3 medium
- 100mm Petri plates containing LP3 medium + 200 mg/L Hygromycin (Hm)
- 100mm Petri plates containing MSK medium + 200 mg/L Hm
- 250 ml culture vessels containing MSO medium + 200mg/L
- Hygromycin stock solution (50 mg/ml in PDS, sterile)

Procedure

1) Harvest and surface sterilise floral tillers of *Lolium perenne* in 5% available chlorine Na-hypochlorite for 15 minutes using a Mason jar (or equivalent) under constant agitation.
2) Rinse tillers with autoclaved ddH_2O.
3) Aseptically dissect floral meristems.
4) Culture meristems on callus induction medium LP5 (16-20 explants per plate) and incubate in the dark for four to six weeks.

5) On the day of transformation transfer embryogenic callus material to high osmotic medium LP3-OS. Arrange approximately 4 cm^2 of calli in the centre of the Petri dish.

6) Incubate calli for 4-6 hours at room temperature.

7) Prepare particles and perform biolistic transformation following the protocol: "Biolistic Transformation of *Lolium perenne* with the Bio-Rad Particle Delivery System (PDS)". Plasmids are co-transformed. One plasmid (pAcH1) contains the hygromycin phosphotransferase gene conferring resistance to the antibiotic hygromycin expressed from the rice actin promoter and the second plasmid contains the genetic construct of interest for transformation. Plasmids are mixed in a one to one ratio at 1μg/μLand simultaneously coated onto the microcarriers.

8) Incubate bombarded calli on high osmotic medium LP3-OS for an additional 12-16 hours (overnight) at 25°C in the dark.

9) Transfer bombarded calli to LP3 medium and incubate for 48 hours at 25°C in the dark

10) Plate calli on selection medium (LP3 + 200 mg/l Hygromycin (Hm)). Incubate at 25°C in the dark on selection medium for two weeks.

11) Transfer all Hm-resistant callus material to regeneration medium MSK + 200 mg/l Hm and incubate for four weeks at 25°C under a 16hour photoperiod.

12) Transfer developed shoots to MS0 + 200 mg/l Hm and incubate for another two to four weeks at 25°C under 16hour photoperiod.

13) Screen by PCR Hm-resistant plants growing on MSO + 200 mg/L Hm.

Microprojectile bombardment of Lolium perenne with the Bio-Rad Particle Delivery System (PDS-1000/He)

Taken from the PDS-100/He manual. These procedures were developed by Sanford *et al.* (1992).

Materials and Solutions

- Bio-Rad Biolistic® PDS-1000/He Particle Delivery System
- Rupture disks (900 PSI)
- Macrocarriers
- Macrocarrier holders
- Microcarriers (1.0 μm)
- Stopping screens
- Autoclaved 1.5 ml eppendorf tubes
- Micropipette tips
- Vortex and microfuge
- Torque wrench tool
- Pen vac
- 70% Ethanol

- Absolute Ethanol
- 2.5 M $CaCl_2$
- 100 mM Spermidine

5 **(A) Microcarrier preparation**

For 120 bombardments using 500 µg per bombardment.

1. In a 1.5 ml microfuge tube, weigh out 60 mg of microparticles.
2. Add 1 ml of 70% ethanol, freshly prepared.
10 3. Vortex on a platform vortexer for 3-5 minutes.
4. Incubate for 15 minutes.
5. Pellet the microparticles by spinning for 5 seconds in a microfuge.
6. Remove the liquid and discard.
7. Repeat the following steps three times:

15 a. Add 1 ml of sterile water

b. Vortex for 1 minute

c. Allow the particles to settle for 1 minute

d. Pellet the microparticles by spinning for 2 seconds in a microfuge.

e. Remove the liquid and discard.

8. Add sterile 50% glycerol to bring the microparticle concentration to 60 mg/ml (assume no loss during preparation).

9. Store the microparticles at room temperature for up to 2 weeks.

(B) Coating DNA onto microcarriers

The following procedure is sufficient for six bombardments; if fewer bombardments are needed, prepare enough microcarriers for three bombardments by reducing all volumes by one half. When removing aliquots of microcarriers, it is important to vortex the tube containing the microcarriers continuously in order to maximise uniform sampling.

1. Vortex the microcarriers prepared in 50% glycerol (60 mg/ml) for 5 minutes on a platform vortexer to resuspend and disrupt agglomerated particles.

2. Remove 50 µl (3 mg) of microcarriers to a 1.5 ml microfuge tube.

3. While vortexing vigorously, add in order:

   5 µl DNA (1 µg/µl)

   50 µl $CaCl_2$ (2.5 M)

   20 µl spermidine (0.1 M)

4. Continue vortexing for 2-3 minutes

5. Allow the microcarriers to settle for 1 minute

6. Pellet the microcarriers by spinning for 2 seconds in a microfuge

7. Remove the liquid and discard

8. Add 140 µl of 70% ethanol without disturbing the pellet

9. Remove the liquid and discard

10. Add 140 µl of 100% ethanol without disturbing the pellet

11. Remove the liquid and discard

12. Add 48 µl of 100% ethanol

13. Gently resuspend the pellet by tapping the side of the tube several times, and then by vortexing at low speed for 2-3 seconds

14. Remove six 6 µl aliquots of microcarriers and transfer them to the centre of a macrocarrier. An effort is made to remove equal amounts (500 µg) of microcarriers each time and to spread them evenly over the central 1 cm of the macrocarrier using the pipette tip. Desiccate immediately.

C) Bombardment procedure

1) Open valve of helium cylinder

2) Adjust helium regulator by turning the helium pressure regulator to 200 PSI above chosen rupture disk (e.g. if a 900 PSI rupture disk will be used, the working pressure has to be adjusted to 1100 PSI)

3) Turn on vacuum pump

4) Place 900psi rupture disk in the rupture disk-retaining cap. Screw on and tighten retaining cap.

5) Place macrocarriers in sterile macrocarrier holder

6) Place stop screen and macrocarrier holder in the launch assembly, tighten screw lid and place below rupture disk-retaining cap. Launch assembly should be set to a Gap distance of !/4 inch and macrocarrier travel distance of 11mm.

7) Place tissue sample at a target distance of 90mm.

8) Turn on main switch of PDS

9) Apply vacuum to 27 inches of Hg

10) Hold vacuum and press "fire" button until shot is performed (automatic)

11) Release "fire" button and vent chamber

12) After shooting close valve of helium cylinder and loosen pressure valve

Table 7. Compositions of the media used

Media component	LP3	LP5	LP3-OS	MSK	MS0

Macro elements (mg/l final concentration) KNO_3 NH_4NO_3 $CaCl_2 \times 2H_2O$ $MgSO_4 \times 2H_2OKH_2PO_4$ KCl	 1900 1650 440 370 170	 1900 1650 440 370 170	 1900 1650 440 370 170	 1900 1650 440 370 170	 1900 1650 440 370 170
Micro elements (mg/l final concentration) Na_2EDTA $FeSO_4 \times 7H_2O$ H_3BO_3 KI $MnSO_4 \times H_2O$ $ZnSO_4 \times 7H_2O$ $CuSO_4 \times 5H_2O$ $Na_2MoO_4 \times 2H_2O$ $CoCl_2 \times 6H_2O$	 37.3 27.8 6.2 0.83 16.9 8.6 0.025 0.25 0.025	 37.3 27.8 6.2 0.83 16.9 8.6 0.025 0.25 0.025	 37.3 27.8 6.2 0.83 16.9 8.6 0.025 0.25 0.025	 37.3 27.8 6.2 0.83 16.9 8.6 0.025 0.25 0.025	 37.3 27.8 6.2 0.83 16.9 8.6 0.025 0.25 0.025
Carbohydrates (g/l final concentration) Maltose D-Mannitol	 30	 30	 30 64	 30	 30
Hormones (mg/l final concentration) 2,4-D Kinetin	 3.0	 5.0	 3.0	 0.2	
Vitamins (mg/l final concentration) Pyridoxine HCl Thiamine HCl Nicotinic acid Myo-Inositol	 0.5 0.1 0.5 100	 0.5 0.1 0.5 100	 0.5 0.1 0.5 100	 0.5 0.1 0.5 100	
Other organics (mg/l final concentration) Glycine	 2	 2	 2	 2	 2

Culture Media

Weights and volumes required of each individual ingredient are specified in Table 7. Adjust media pH to 5.8 with KOH. The addition of a solidifying agent is required. Use agarose (for LP3, LP5 and LP3-OS) and 0.8% (w/v) Agar for MS0 and MSK prior to sterilising. Media LP3, LP5 and MSK are modified from Murashige and Skoog (1962).

Expression of chimeric genes in Corn Cells

A chimeric gene comprising a lolitrem cDNA encoding in sense orientation with respect to the promoter that is located 5' to the cDNA fragment, and a terminator 3' to the cDNA fragment, can be constructed. The cDNA fragment of this gene may be generated by polymerase chain reaction (PCR) of the cDNA clone using appropriate oligonucleotide primers. Cloning sites can be incorporated into the oligonucleotides to provide proper orientation of the DNA fragment when inserted into the digested vector as described below. Amplification is then performed in a standard PCR reaction. The amplified DNA is then digested with restriction enzymes and fractionated on an agarose gel. The appropriate band can be isolated from the gel and combined with Vector and insert DNA can be ligated at 15°C overnight, essentially as described (Sambrook). The ligated DNA may then be used to transform *E.Coli* XL1-Blue (Epicurian Coli XL-1 Blue™, Stratagent). Bacterial transformants can be screened by restriction enzyme digestion of plasmid DNA and limited nucleotide sequence analysis using the dideoxy chain termination method (Sequenase™ DNA sequencing Kit; US Biochemical). The resulting plasmid construct would comprise a chimeric gene encoding in the 5' to 3' direction promoter, a cDNA encoding and the 3' region containg a terminator.

The chimeric gene described above can then be introduced into cells by the following procedure. Immature corn embryos can be dissected from developing caryopses derived from crosses of the inbred corn lines. The embryos are isolated 10 to 11 days after pollination when theya re 1.0 to 1.5mm long. The embryos are then placed with the axis-side facing down and in contact with agarose-solidified N6 medium (Chu et al. (1975) *Sci. Sin Peking* 18:659-668). The embryos are kept in the dark at 27°C. Friable embryogenic callus consisting of undifferentiated masses of cells with somatic proembryoids and embryoids borne on suspensor structures proliferates from the scutellum of these immature embryos. The embryogenic callus isolated from the

primary explant can be cultured on N6 medium and sub-cultured on this medium every 2 to 3 weeks.

The particle bombardment method (Klein et al. (1987) *Nature* 327:70-73) may be used to transfer genes to the callus culture cells. According to this method, gold particles (1um in diatmeter) are coated with DNA using the following technique. Ten ug of plasmid DNAs are added to 50 uL of a suspension of gold particles (60mg per mL). Calcium chloride (50 uL of a 2.5 M solution) and spermidine free base (20 uL of a 1.0M solution) are added to the particles. The suspension is vortexed during the addition of these solutions. After 10 minutes, the tubes are briefly centrifuged (5 sec at 15,000 rpm) and the supernatant removed. The particles are resuspended in 200 uL of absolute ethanol, centrifuged again and the supernatant removed. The ethanol rinse is performed again and the particles resuspended in a final volume of 30 uL of ethanol. An aliquot (5 uL) of the DNA-coated gold particles can be placed in the centre of a Kapton™ flying disc (Bio-Rad Labs). The particles are then accelerated into the corn tissue with a Biolistic™ PDS-1000/He (Bio-Rad Instruments Hercules CA), using a helium pressure of 1000 psi, a gap distance of 0.5 cm and a flying distance of 1.0 cm.

For bombardment, the embryogenic tissue is placed on filter paper over agarose-solidified N6 medium. The tissue is arranged as a thin lawn and covered a circular area of about 5 cm in diameter. The Petri dish containing the tissue can be placed in the chamber of the PDS-1000/He approximately 8 cm from the stopping screen. The air in the chamber is then evacuated to a vacuum of 28 inches of Hg. The macrocarrier is accelerated with a helium shock wave using a rupture membrane that bursts when the He pressure in the shock tube reaches 1000 psi.

Seven days after bombardement the tissue can be transferred to N6 medium that contains a selection. The tissue continues to grow slowly on this medium. After an additional 2 weeks the tissue can be transferred to fresh N6 medium containing the selection. After 6 weeks, areas of about 1 cm in diameter of actively growing callus can be identified on some of the plates containing the selective medium. These calluses may continue to grow when sub-cultured on the selective medium.

Plants can be regenerated from the transgenic callus by first transferring clusters of tissue to N6 medium supplemented with 0.2 mg per liter of 2,4-D. After two weeks the tissue can be transferred to regeneration medium (Fromm et cl., (1990) *BioTechnology* 8:833-839.

Aspects of the present invention have been described by way of example only and it should be appreciated that modifications and additions may be made thereto without departing from the scope thereof.

REFERENCES

Acklin W, Weibel F, Arigoni D (1977) Zur biosynthese von paspalin und verwandten metaboliten aus *Claviceps paspali*. Chimia 31:63

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403-410

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucl Acids Res 25:3389-3402

Barbato C, Calissano M, Pickford A, Romano N, Sandmann G, Macino G (1996) Mild RIP - an alternative method for *in vivo* mutagenesisof the *albino-3* gene in *Neurospora crassa*. Mol Gen Genet 252:353-361

Bush LP, Wilkinson HH, Schardl CL (1997) Bioprotective alkaloids of grass-fungal endophyte symbioses. Plant Physiol 114:1-7

Carroll AM, Sweigard JA, Valent B (1994) Improved vectors for selecting resistance to hygromycin. Fungal Genet News 22

Chen AP, Kroon PA, Poulter CD (1994) Isoprenyl diphosphate synthases: protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure. Protein Sci 3:600-607

Cole RJ, Cox RH (1981) Handbook of toxic fungal metabolites. Academic Press, London

de Jesus AE, Gorst-Allman CP, Steyn PS, van Heerden FR, Vleggar R, Wessels PL, Hull WE (1983) Tremorogenic mycotoxins from *Penicillium crustosum*. Biosynthesis of Penitrem A. J. Chem. Soc. Perkin Trans. 1863-1868

Eggink G, Engel H, Vriend G, Terpstra P, Witholt B (1990) Rubredoxin reductase of *Pseudomonas oleovorans*: structural relationship to other flavoprotein oxidoreductases based on one NAD and two FAD fingerprints. J Mol Biol 212:135-142

Fletcher LR (1999) "Non-toxic" endophytes in ryegrass and their effect on livestock health and production. In: Woodfield DR, Matthew C (eds) Ryegrass endophyte: an essential New Zealand symbiosis. Napier, New Zealand, pp. 133-139

Gatenby WA, Munday-Finch SC, Wilkins AL, Miles CO (1999) Terpendole M, a novel indole-diterpenoid isolated from *Lolium perenne* infected with the endophytic fungus *Neotyphodium lolii*. J Agric Food Chem 47:1092-1097

Gloer JB (1995) Antiinsectan natural products from fungal sclerotia. Acc Chem Res 28:343-350

Graham-Lorence SE, Peterson JA (1996) Structural alignments of P450s and extrapolations to the unknown. Methods Enzymol 272:315-326

Higgins D, Thompson JN, Gibson TJ, Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,position-specific gap penalties and weight matrix choice. Nucl Acids Res 22:4673-4680

Huang X-H, Tomoda H, Nishida H, Masuma R, Omura S (1995) Terpendoles, novel ACAT inhibitors produced by *Albophoma yamanashiensis*. I. Production, isolation and biological properties. J Antibiotics 48:1-4

Keller NP, Hohn TM (1997) Metabolic pathway gene clusters in filamentous fungi. Fungal Genet Biol 21:17-29

Knaus H-G, McManus OB, Lee SH, Schmalhofer WA, Garcia-Calvo M, Helms LMH, Sanchez M, Giangiacomo K, Reuben JP, Smith AB, Kaczorowski GJ, Garcia ML (1994) Tremorgenic indole alkaloids potently inhibit smooth muscle high-conductance calcium-activated channels. Biochemistry 33:5819-5828

Laakso JA, Gloer JB, Wicklow DT, Dowd PF (1992) Sulpinines A-C and secopenitrem B: new antiinsectan metabolites from the sclerotia of *Aspergillus sulphurous*. J Org Chem 57:2066-2071

Latch GCM, Christensen MJ (1985) Artificial infection of grasses with endophytes. Annals of Applied Biology 107:17-24

Laws I, Mantle PG (1989) Experimental constraints in the study of the biosynthesis of indole alkaloids in fungi. J Gen Microbiol 135:2679-2692

Li C, Gloer JB, Wicklow DT, Dowd PF (2002) Thiersinines A and B: novel antiinsectan indole diterpenoids from a new fungicolous *Penicillium* species (NRRL 28147). Org Lett 4:3095-3098Mantle PG (1987) Secondary metabolites of *Penicillium* and *Acremonium*. In: Peberdy JF (ed), *Penicillium* and *Acremonium*. Plenum Press, New York, pp. 161-243

Mantle PG, Weedon CM (1994) Biosynthesis and transformation of tremorgenic indole-diterpenoids by *Penicillium paxilli* and *Acremonium lolii*. Phytochemistry 36:1209-1217

McMillan LK, Carr RL, Young CA, Astin JW, Lowe RGT, Parker EJ, Jameson GB, Finch SC, Miles CO, McManus OB, Schmalhofer WA, Garcia ML, Kaczorowski GJ, Goetz MA, Tkacz JS, Scott B (2003) Molecular analysis of two cytochrome P450 monooxygenase genes required for paxilline biosynthesis in *Penicillium paxilli* and effects of paxilline intermediates on mammalian maxi-K ion channels. Mol Gen Genom 270:9-23

Mende K, Homann V, Tudzynski B (1997) The geranylgeranyl diphosphate synthase gene of *Gibberella fujikuroi*: isolation and expression. Mol Gen Genet 255:96-105

Munday-Finch SC, Wilkins AL, Miles CO (1996) Isolation of paspaline B, an indole-diterpenoid from *Penicillium paxilli*. Phytochemisty 41:327-332

Ondeyka JG, Helms GL, Hensens OD, Goetz MA, Zink DL, Tsipouras A, Shoop WL, Slayton L, Dombrowski AW, Polishook JD, Ostlind DA, Tsou NN, Ball RG, Singh SB (1997) Nodulisporic acid A, a novel and potent insecticide from a *Nodulisporum* sp. Isolation, structure determination and chemical transformations. J Chem Soc Chem Commun 119:8809-8816

Parker EJ, Scott DB (2004) Indole-diterpene biosynthesis in ascomycetous fungi. In: An Z (ed), Handbook of Industrial Mycology. Marcel Dekker, New York,

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85:2444-2448

Popay AJ, Hume DE, Baltus JG, Latch GCM, Tapper BA, Lyons TB, Cooper BM, Pennell CG, Eerens JPJ, Marshall SL (1999) Field performance of perennial ryegrass (*Lolium perenne*) infected with toxin-free fungal endophytes (*Neotyphodium spp.*). In: Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York

Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463-5467

Schardl CL (2001) *Epichloë festucae* and related mutualistic symbionts of grasses. Fungal Genet Biol 33:69-82

Schardl CL (1996) *Epichloë* species: fungal symbionts of grasses. Ann Rev Phytopathol 34:109-130

Scott B (2001) *Epichloë* endophytes: symbionts of grasses. Curr Opin Microbiol 4:393-398

Scott DB, Jameson GB, Parker EJ (2003) Isoprenoids: gene clusters and biosynthetic puzzles. In: Tkacz JS, Lange L (eds), Advances in fungal biotechnology for industry, agriculture and medicine. Kluwer Academic/Plenum Publishers, New York,

Selala MI, Laekeman GM, Loenders B, Masuka A, Herman AG, Schepens P (1991) *In vitro* effects of tremorgenic mycotoxins. J Nat Products 54:207-212

Steyn PS, Vleggaar R (1985) Tremorgenic mycotoxins. Prog Chem Organic Natural Products 48:1-80

Tudzynski B, Hölter K (1998) Gibberellin biosynthetic pathway in *Gibberella fujikuroi*: evidence for a gene cluster. Fungal Genet Biol 25:157-170

Vallon O (2000) New sequence motifs in flavoproteins: evidence for common ancestry and tools to predict structure. Proteins: Structure, Function and Genetics 38:95-114

Wang J, Machado C, Panaccione DG, Tsai H-F, Schardl CL (2004) The determinant step in ergot alkaloid biosynthesis by an endophyte of perennial ryegrass. Fungal Genet Biol

Wierenga RK, Terpstra P, Hol WGJ (1986) Predictions of the occurrence of the ADP-binding bab-fold in proteins using an amino acid sequence fingerprint. J Mol Biol 187:101-107

Woodfield DR, Matthew C (eds) Ryegrass endophyte: an essential New Zealand symbiosis. Napier, New Zealand, pp. 113-122

Yao Y, Peter AB, Baur R, Sigel E (1989) The tremorgen aflatrem is a positive allosteric modulator of the gamma-aminobutyric acid receptor channel in *Xenopus* oocytes. Mol Pharmacol 35:319-323

Young CA, McMillan L, Telfer E, Scott B (2001) Molecular cloning and genetic analysis of an indole-diterpene gene cluster from *Penicillium paxilli*. Mol Microbiol 39:754-764

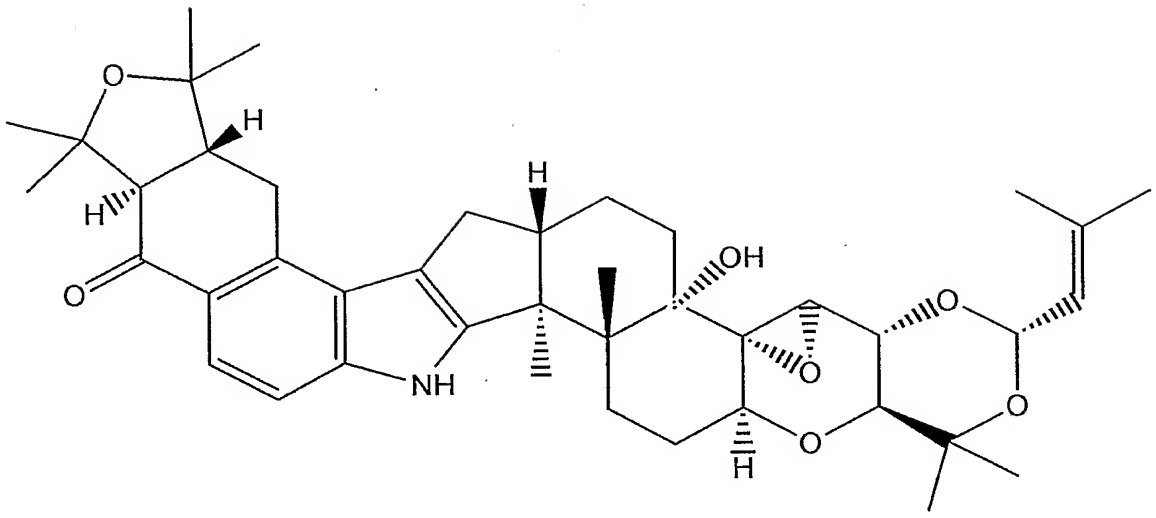
Young C, Itoh Y, Johnson R, Garthwaite I, Miles CO, Munday-Finch SC, Scott B (1998) Paxilline-negative mutants of *Penicillium paxilli* generated by heterologous and homologous plasmid integration. Curr Genet 33:368-377

AgResearch Limited

by their attorneys

JAMES & WELLS

Figure 1.



Lolitrem B

Figure 2

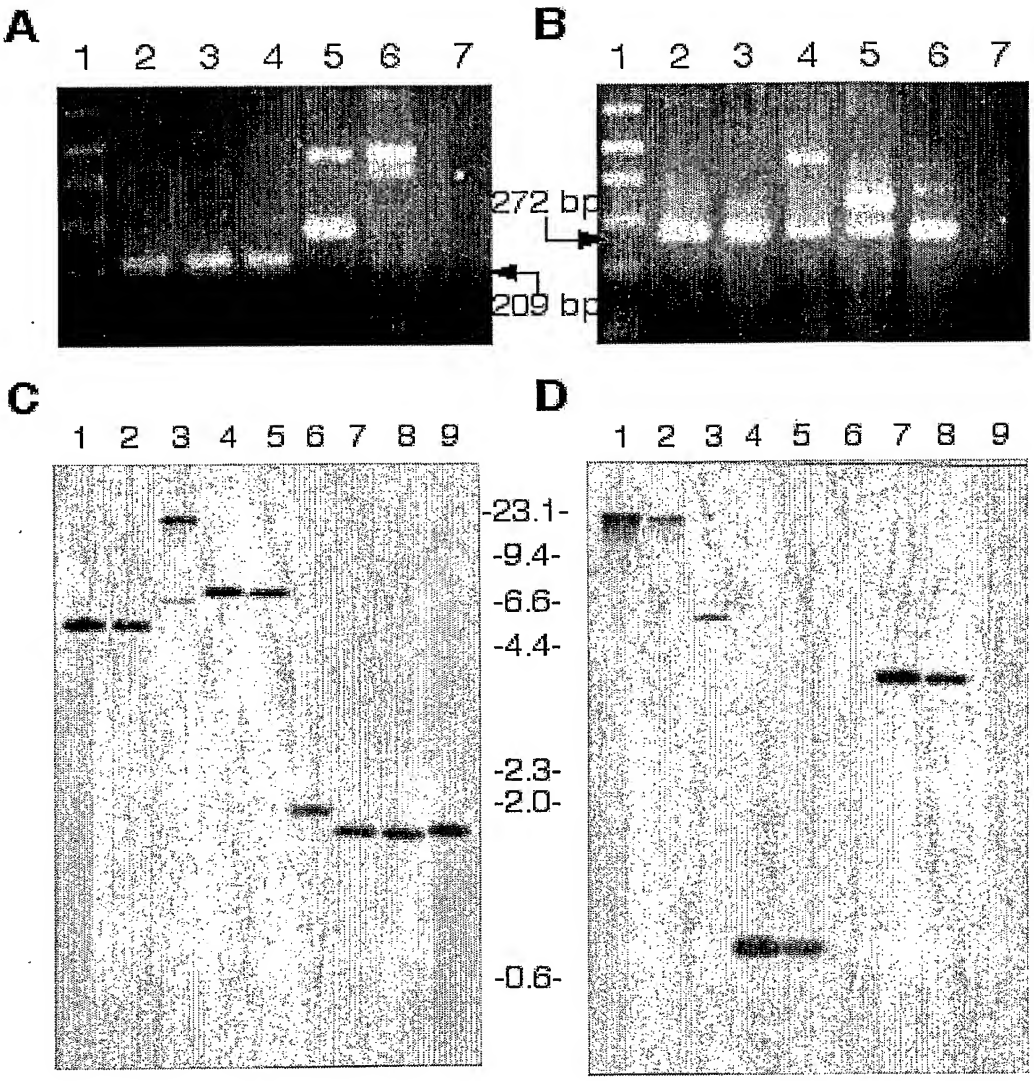

A
1 2 3 4 5 6 7
B
1 2 3 4 5 6 7
272 bp
209 bp
C
1 2 3 4 5 6 7 8 9
-23.1-
-9.4-
-6.6-
-4.4-
-2.3-
-2.0-
-0.6-
D
1 2 3 4 5 6 7 8 9

Figure 3

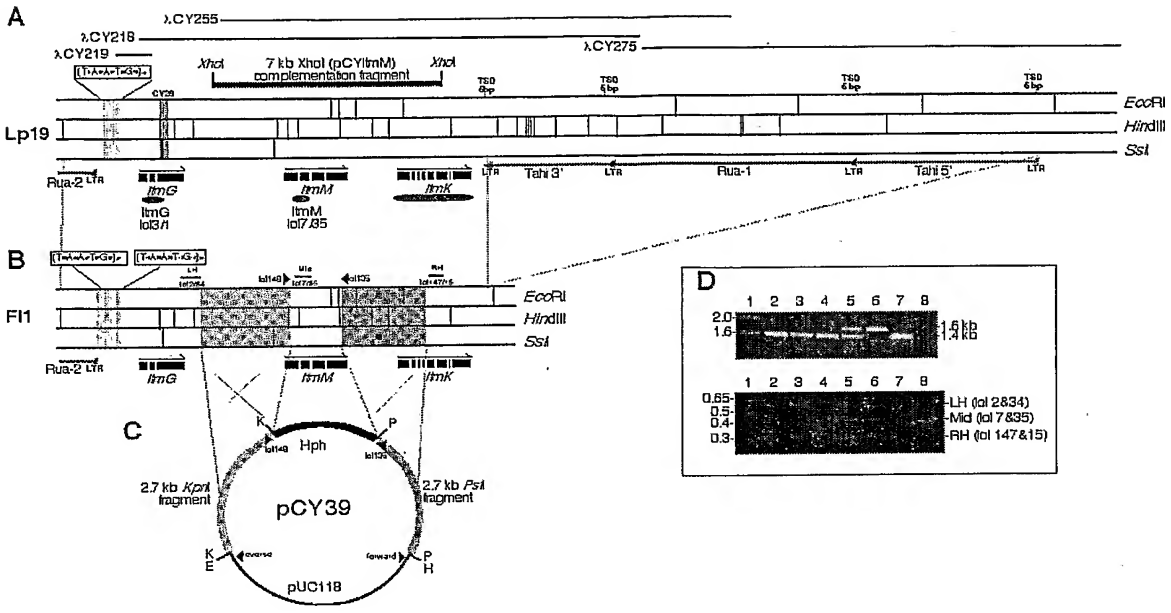

A
λCY255
λCY218
λCY219
λCY275
XhoI
7 kb XhoI (pCY1tmM) complementation fragment
XhoI
TSD 6bp
EcoRI
HindIII
SstI
Lp19
Rua-2 LTR
ltmG
ltmM
ltmK
ltmG lol3A
ltmM lol7&35
LTR
Tahi 3'
Rua-1
Tahi 5'
B
Fl1
C
K
P
Hph
2.7 kb KpnI fragment
2.7 kb PstI fragment
pCY39
pUC118
D
1 2 3 4 5 6 7 8
2.0
1.6
1.5 kb
1.4 kb
0.65
0.5
0.4
0.3
LH (lol 2&34)
Mid (lol 7&35)
RH (lol 147&15)

Figure 4.

```
   1  [redacted]ACGATGGCTGCCAATGACTTTCCATTTCAATGCCAGGAGAAGAAATC
  51  ATATTCTCAGCCAAGTCTAGTCTACTGCAATGGTAACATTGCGGAGACGT
 101  ATCTCGAAGAAAAG[redacted]
 151  [redacted]GTTTTGACAGCGCCGTTGGATTATTTGCGTGCCTT
 201  ACCTAGCAAAGATATTCGCAGTGGACTGACCGACGCCATTAATGAGTTCC
 251  TGCGTGTCCCAGAGGAAAAGGTTCTTGTCATAAAGCGTATAATTGATCTT
 301  CTTCACAATGCATCCTTACT[redacted]
 351  [redacted]CATTGATGATATCCAGGATTCATCCA
 401  AACTGCGACGTGGAGTCCCTGTAGCCCACCACATATTTGGAATCGCACAA
 451  ACAATAAATTCGGCCAATCTAGCGTATTTCATTGCCCAGAGAGAGCTTGA
 501  GAAGCTTACGAATCCTCGAGCATTTGCTATATATAATGAGGAGCTAATCA
 551  ATCTGCATCGTGGTCAGGGTATGGAGCTCCATTGGAGAGAATCGCTCCAT
 601  TGCCCTACCGAAGATGAGTATCTGCGAATGATCCAAAAGAAGACAGGCGG
 651  TCTGTTCCGATTGGCAATCAGACTGCTGCAAGGCGAAAGCGCTAGCGATG
 701  ACGATTATGTCTCACTTATTGATACTCTCGGAACCCTGTTCCAGATTCGA
 751  GATGACTATCAAAACTTACAGAGTGATATATATTCTAAGAACAAAGGCTA
 801  CTGTGAGGATTTAACAGAGGGCAAATTCTCGTATCCGGTCATCCATAGTA
 851  TTCGGTCGCGACCAGGAGATGTTCGATTAATCAATATTTTGAAACAGCGT
 901  AGTGAAGATGTTATGGTGAAGCAATACGCGGTGCAACATATCGAATCTAC
 951  AGGAAGCTTCGCATTCTGTCAAAATAAAATTCAATCTTTGGTGGAGCAAG
1001  CAAGAGAGCAATTGGCGGCTCTAGAAAATAGCAGTTCATGTGGAGGCCCC
1051  GTTCGCGACATCCTTGACAAGTTAGCAATAAAACCACGGGCAAATATAGA
1101  AGTAGAG[redacted]
```

Figure 5.

```
   1  MTMAANDFPF QCQEKKSYSQ PSLVYCNGNI AETYLEEKVL TAPLDYLRAL
  51  PSKDIRSGLT DAINEFLRVP EEKVLVIKRI IDLLHNASLL IDDIQDSSKL
 101  RRGVPVAHHI FGIAQTINSA NLAYFIAQRE LEKLTNPRAF AIYNEELINL
 151  HRGQGMELHW RESLHCPTED EYLRMIQKKT GGLFRLAIRL LQGESASDDD
 201  YVSLIDTLGT LFQIRDDYQN LQSDIYSKNK GYCEDLTEGK FSYPVIHSIR
 251  SRPGDVRLIN ILKQRSEDVM VKQYAVQHIE STGSFAFCQN KIQSLVEQAR
 301  EQLAALENSS SCGGPVRDIL DKLAIKPRAN IEVE
```

Figure 6.

```
   1  [redacted]ACTAGCGACTTCAAGGTAATAATCGTGGGAGGATCAGTGGCTGGGCT
  51  TTCACTAGCCCACTGCTTAGAAAAAATCGGTGTTTCTTTCATGGTTCTAG
 101  AGAAGGGTAATCAAATAGCTCCCCAACTCGGTGCCTCAATTGGCATTTTG
 151  CCAAATGGTGGACGTATTCTTGATCAACTGGGCATCTTCCATAGCATCGA
 201  GGATGAAATCGAACCTCTAGAATCTGCTATGATGAGATACCCGGATGGTT
 251  TCTCTTTCAAAAGTCAATATCCCCAAGCTTTGCATACTAG[redacted]
 301  [redacted]
 351  [redacted]TTTTGGTTATCCCGTGGCTTTCCTTGAGAGGCAAAGGTTTCTTC
 401  AGATACTTTATGATAAACTCAAGAGCAAAGACTGCGTTTTTACAAACAAG
 451  CGGGTAGTCAGTATTGCAAGTGGCCAAGACAAAGTCACAGCAAAGACTTC
 501  AGATGGCGCTAAGTACTTAGCAGATATCGTGATCGGTGCTGACGGGGTCC
 551  ACAGCATCGTCAGGTCAGAGATTTGGAGGCATTTGAAGGAAAACTCTCAA
 601  ATATCAGTATTAGAGGCACCGAACGCAA[redacted]
 651  [redacted]GTATTAAGCATGATTA
 701  TTCATGCATTTACGGAATTTCTTTAAACGTTCCCCAGATCATCCTAGGAA
 751  TACAGTTAAACTGTTTAGATGACGGAGTGTCAATACACTTGTTTACGGGT
 801  AAACAATCCAAATTATTTTGGTTTGTTATCATCAAAACGCCTCAGGCTAG
 851  CTTTGCTAAAGTAGAGATTGACAATACACATACAGCAAGGTGTATCTGCG
 901  AAGGACTGAGGACGAAAAAGGTTTCAGATACCTTATGTTTTGAAGATGTA
 951  TGGTCAAGATGCACCATATTCAAGATGACGCCTCTTGAGGAAGGGGTGTT
1001  TAAGCATTGGAACTATGGCCGCTTAGCATGTATTGGTGATGCTATCCGCA
1051  AG[redacted]
1101  [redacted]
1151  [redacted]ATGGCCCCAAATAATGGGCAAGGAGCAAATATGGCGATAGAG
1201  GACGCTTGCAGTCTCGCAAACATCCTCCAGAAAAAGATATCACATGGTTC
1251  GATTCGAGACCAAGATATCAATTCAATGTTTCAGGAATTCTCTATGGCTC
1301  AACGGGCTCGCACGGAGAGCGTCTGCGCGCAGTCGGAGTTTCTAGTCCGC
1351  ATGCATGCGAATCAAGGTATTGGAAGAAGACTTCTTGGGCGGTACCTTAT
1401  TCCTTTCCTGTATGACGCACCTGCTGGTTTATCTGGATTTTCTATAAGTG
1451  GCGCAACAAGAATAGAGTTCATAGACTTGCCCACTAGATCTCTTAGGGGA
1501  GCGTGGGGAAAGTCATGGAGAGGGTCATGGGAATTCATCCTACAAAGCTT
1551  GGTCTATTTGCGACCCAAGTTTAGGATAGTTTATGCCTTGTATCTCGTTG
1601  CAGCTGCAGCTTTTATCTTGTATTGTCTTAGCAGTCTCTTCCCG[redacted]
```

Figure 7.

```
   1  MTSDFKVIIV GGSVAGLSLA HCLEKIGVSF MVLEKGNQIA PQLGASIGIL
  51  PNGGRILDQL GIFHSIEDEI EPLESAMMRY PDGFSFKSQY PQALHTSFGY
 101  PVAFLERQRF LQILYDKLKS KDCVFTNKRV VSIASGQDKV TAKTSDGAKY
 151  LADIVIGADG VHSIVRSEIW RHLKENSQIS VLEAPNASIK HDYSCIYGIS
 201  LNVPQIILGI QLNCLDDGVS IHLFTGKQSK LFWFVIIKTP QASFAKVEID
 251  NTHTARCICE GLRTKKVSDT LCFEDVWSRC TIFKMTPLEE GVFKHWNYGR
 301  LACIGDAIRK MAPNNGQGAN MAIEDACSLA NILQKKISHG SIRDQDINSM
 351  FQEFSMAQRA RTESVCAQSE FLVRMHANQG IGRRLLGRYL IPFLYDAPAG
 401  LSGFSISGAT RIEFIDLPTR SLRGAWGKSW RGSWEFILQS LVYLRPKFRI
 451  VYALYLVAAA AFILYCLSSL FP
```

Figure 8.

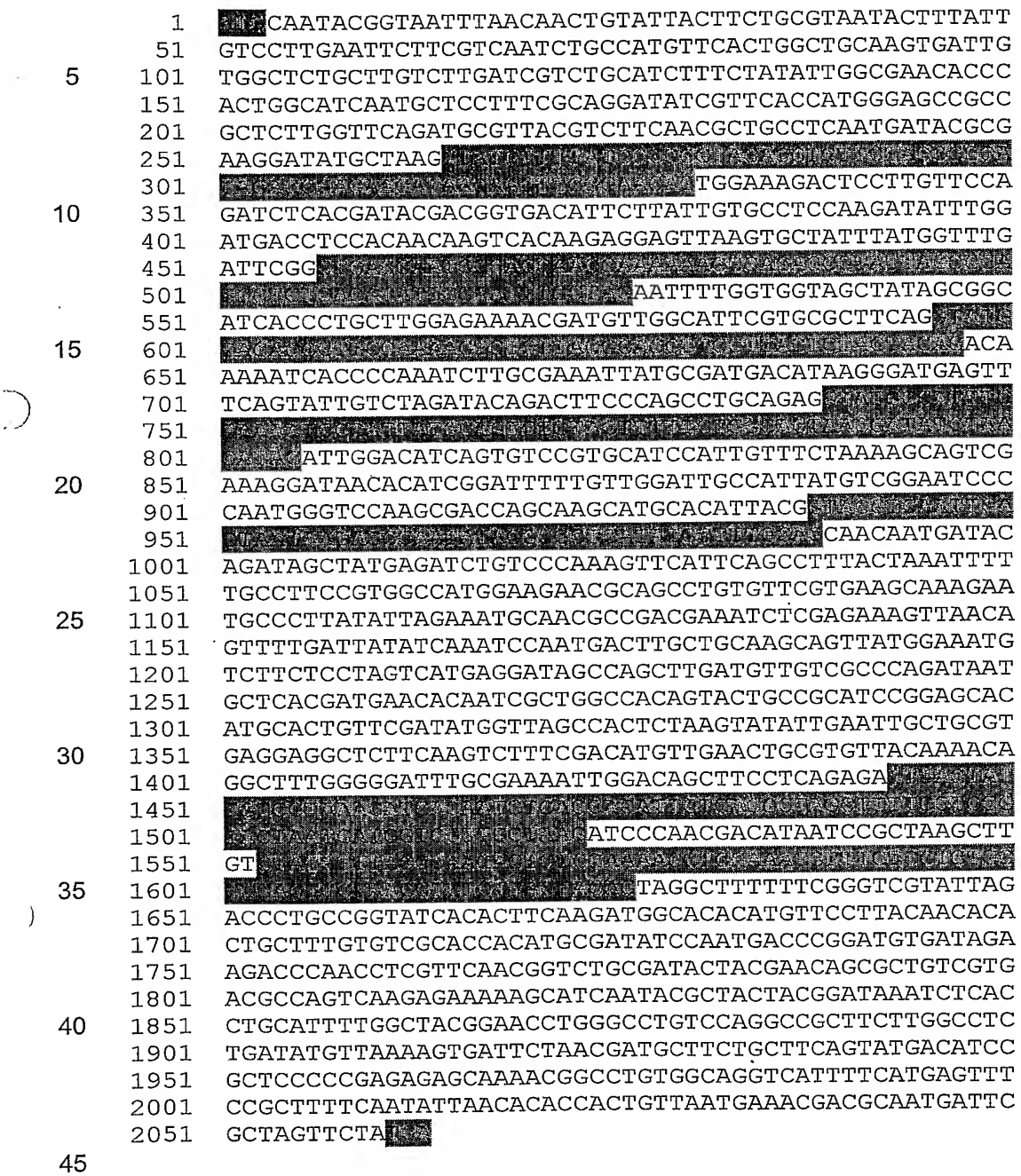

1 CAATACGGTAATTTAACAACTGTATTACTTCTGCGTAATACTTTATT
51 GTCCTTGAATTCTTCGTCAATCTGCCATGTTCACTGGCTGCAAGTGATTG
101 TGGCTCTGCTTGTCTTGATCGTCTGCATCTTTCTATATTGGCGAACACCC
151 ACTGGCATCAATGCTCCTTTCGCAGGATATCGTTCACCATGGGAGCCGCC
201 GCTCTTGGTTCAGATGCGTTACGTCTTCAACGCTGCCTCAATGATACGCG
251 AAGGATATGCTAAG
301 TGGAAAGACTCCTTGTTCCA
351 GATCTCACGATACGACGGTGACATTCTTATTGTGCCTCCAAGATATTTGG
401 ATGACCTCCACAACAAGTCACAAGAGGAGTTAAGTGCTATTTATGGTTTG
451 ATTCGG
501 AATTTTGGTGGTAGCTATAGCGGC
551 ATCACCCTGCTTGGAGAAAACGATGTTGGCATTCGTGCGCTTCAG
601 ACA
651 AAAATCACCCCAAATCTTGCGAAATTATGCGATGACATAAGGGATGAGTT
701 TCAGTATTGTCTAGATACAGACTTCCCAGCCTGCAGAG
751
801 ATTGGACATCAGTGTCCGTGCATCCATTGTTTCTAAAAGCAGTCG
851 AAAGGATAACACATCGGATTTTTGTTGGATTGCCATTATGTCGGAATCCC
901 CAATGGGTCCAAGCGACCAGCAAGCATGCACATTACG
951 CAACAATGATAC
1001 AGATAGCTATGAGATCTGTCCCAAAGTTCATTCAGCCTTTACTAAATTTT
1051 TGCCTTCCGTGGCCATGGAAGAACGCAGCCTGTGTTCGTGAAGCAAAGAA
1101 TGCCCTTATATTAGAAATGCAACGCCGACGAAATCTCGAGAAAGTTAACA
1151 GTTTTGATTATATCAAATCCAATGACTTGCTGCAAGCAGTTATGGAAATG
1201 TCTTCTCCTAGTCATGAGGATAGCCAGCTTGATGTTGTCGCCCAGATAAT
1251 GCTCACGATGAACACAATCGCTGGCCACAGTACTGCCGCATCCGGAGCAC
1301 ATGCACTGTTCGATATGGTTAGCCACTCTAAGTATATTGAATTGCTGCGT
1351 GAGGAGGCTCTTCAAGTCTTTCGACATGTTGAACTGCGTGTTACAAAACA
1401 GGCTTTGGGGGATTTGCGAAAATTGGACAGCTTCCTCAGAGA
1451
1501 ATCCCAACGACATAATCCGCTAAGCTT
1551 GT
1601 TAGGCTTTTTTCGGGTCGTATTAG
1651 ACCCTGCCGGTATCACACTTCAAGATGGCACACATGTTCCTTACAACACA
1701 CTGCTTTGTGTCGCACCACATGCGATATCCAATGACCCGGATGTGATAGA
1751 AGACCCAACCTCGTTCAACGGTCTGCGATACTACGAACAGCGCTGTCGTG
1801 ACGCCAGTCAAGAGAAAAAGCATCAATACGCTACTACGGATAAATCTCAC
1851 CTGCATTTTGGCTACGGAACCTGGGCCTGTCCAGGCCGCTTCTTGGCCTC
1901 TGATATGTTAAAAGTGATTCTAACGATGCTTCTGCTTCAGTATGACATCC
1951 GCTCCCCCGAGAGAGCAAAACGGCCTGTGGCAGGTCATTTTCATGAGTTT
2001 CCGCTTTTCAATATTAACACACCACTGTTAATGAAACGACGCAATGATTC
2051 GCTAGTTCTA

Figure 9.

```
   1  MQYGNLTTVL LLRNTLLSLN SSSICHVHWL QVIVALLVLI VCIFLYWRTP
  51  TGINAPFAGY RSPWEPPLLV QMRYVFNAAS MIREGYAKWK DSLFQISRYD
 101  GDILIVPPRY LDDLHNKSQE ELSAIYGLIR NFGGSYSGIT LLGENDVGIR
 151  ALQTKITPNL AKLCDDIRDE FQYCLDTDFP ACRDWTSVSV HPLFLKAVER
 201  ITHRIFVGLP LCRNPQWVQA TSKHAHYATM IQIAMRSVPK FIQPLLNFCL
 251  PWPWKNAACV REAKNALILE MQRRRNLEKV NSFDYIKSND LLQAVMEMSS
 301  PSHEDSQLDV VAQIMLTMNT IAGHSTAASG AHALFDMVSH SKYIELLREE
 351  ALQVFRHVEL RVTKQALGDL RKLDSFLRES QRHNPLSLLG FFRVVLDPAG
 401  ITLQDGTHVP YNTLLCVAPH AISNDPDVIE DPTSFNGLRY YEQRCRDASQ
 451  EKKHQYATTD KSHLHFGYGT WACPGRFLAS DMLKVILTML LLQYDIRSPE
 501  RAKRPVAGHF HEFPLFNINT PLLMKRRNDS LVL
```

Figure 10.

```
      1  AATGGACTAGAAAGTACATTTGTTATACAGTGCTATCTCCTTAGGCTCAG
     51  TCTACCTTGTGGGTCAGTGCAGGCCCCACAGGCCCCCTGCCACAAGGTTA
    101  GTAACCGCGCAAGCACGCGAAAGTGTAGCGTAGTAAATTATATAGGAAAA
    151  ATTAGCAGTATATTAATTATTAGCCTATCTATATATAAGTAAATATACCT
    201  TTAATTCACTTCTATTTAATTGGATATAGACCCTAGTTAACGTGACTTCA
    251  CAAGGTGAACTAAGTCCAAGAAGATAGAGGTAATTGCAGTGAGATCCACA
    301  GGTCTTGTCAGGGGACGGCAATGTATGCATATATCGTGAAATCAATGCTA
    351  GCGGCATTGAATCAATGACTTCTGTAGCTAGCGATAATAGCAGCGATAGA
    401  AGCCTCTAGAATCTATATAGACAGTATTAAGTAAACTCTCCACCTGTATC
    451  CACAGCTAACTTACATACACCTAGCCCTGTCTTGAGTGCTTTTAGAAGAC
    501  TATGCTAACTTAGATCACACCCTAAGTGCCAATGTCTCCCAATTAGCCGC
    551  GAAGAGAGAACTTATCGCAAGGAAGTGATAAGGCTATAACATCCAACAGG
    601  TTACTTAAAGACAACAGGCTAGGAATCAATTATAGTAGCAATCAAAACTA
    651  GATCCTGTATTCTATAACAAGAAGTTAAATCCCCCCTAGACTATCTGTCT
    701  ATCTTTAGTTATACTTTGGTTTTGCTTTGTTGTCTTATGCCTACATTCCT
    751  AAAAGATCTTT[redacted]ACGATGGCTGCCAATGACTTTCCATTTCAATGCCAG
    801  GAGAAGAAATCATATTCTCAGCCAAGTCTAGTCTACTGCAATGGTAACAT
    851  TGCGGAGACGTATCTCGAAGAAAAG[redacted]
    901  [redacted]GTTTTGACAGCGCCGTTGGATTAT
    951  TTGCGTGCCTTACCTAGCAAAGATATTCGCAGTGGACTGACCGACGCCAT
   1001  TAATGAGTTCCTGCGTGTCCCAGAGGAAAAGGTTCTTGTCATAAAGCGTA
   1051  TAATTGATCTTCTTCACAATGCATCCTTACT[redacted]
   1101  [redacted]CATTGATGATATCCA
   1151  GGATTCATCCAAACTGCGACGTGGAGTCCCTGTAGCCCACCACATATTTG
   1201  GAATCGCACAAACAATAAATTCGGCCAATCTAGCGTATTTCATTGCCCAG
   1251  AGAGAGCTTGAGAAGCTTACGAATCCTCGAGCATTTGCTATATATAATGA
   1301  GGAGCTAATCAATCTGCATCGTGGTCAGGGTATGGAGCTCCATTGGAGAG
   1351  AATCGCTCCATTGCCCTACCGAAGATGAGTATCTGCGAATGATCCAAAAG
   1401  AAGACAGGCGGTCTGTTCCGATTGGCAATCAGACTGCTGCAAGGCGAAAG
   1451  CGCTAGCGATGACGATTATGTCTCACTTATTGATACTCTCGGAACCCTGT
   1501  TCCAGATTCGAGATGACTATCAAAACTTACAGAGTGATATATATTCTAAG
   1551  AACAAAGGCTACTGTGAGGATTTAACAGAGGGCAAATTCTCGTATCCGGT
   1601  CATCCATAGTATTCGGTCGCGACCAGGAGATGTTCGATTAATCAATATTT
   1651  TGAAACAGCGTAGTGAAGATGTTATGGTGAAGCAATACGCGGTGCAACAT
   1701  ATCGAATCTACAGGAAGCTTCGCATTCTGTCAAAATAAAATTCAATCTTT
   1751  GGTGGAGCAAGCAAGAGAGCAATTGGCGGCTCTAGAAAATAGCAGTTCAT
   1801  GTGGAGGCCCCGTTCGCGACATCCTTGACAAGTTAGCAATAAAACCACGG
   1851  GCAAATATAGAAGTAGAG[redacted]TTGACATTAAGAACATTGCGATAAAAGAC
   1901  ACTTTTACTATACTCGACTAGTTTTAAAACTATGTGTGAGATTAAGACGT
   1951  CTTCAGGTACTCAAAGTGTGGAAGTATGTCACGCAGAAAAGAGCTAACAT
   2001  TGCTCTCAGCTTCCTCACTATTTAGTTTCACCAAGAGCATCCTTCATAGA
   2051  GACATTTGCGGCTGTGATTTTCGTTTACGTCATGTTGTTAAACATTGTTG
   2101  TATGGTATCTTTGCTTAGGAGTAGACATCCATTTTCTCTCACTCTACTCT
   2151  TAGAGATCGTCAAGTGTCACATACATTTCTGAGAACTAGGACTTTGCATA
   2201  GAATATGCATCGGTTAGGTGTTTGCGTAGAGAGTACGTGTGTCTGAGGTT
   2251  AGCCATTGCGCTTCGTTTGCGGTTTACAATGGGGCAAGGCTTAAAGCTTT
   2301  TTAAAGCCACGGTGACCACTACTGCAGGTGCATTCTTTTTTTAGTCGTAA
   2351  AACACTAAGTTTTTTTTACTAGTTATAATAGACTTTTCTTTCCTTCTTCC
   2401  CTTCTCGTAGATAAACCCAATTGAAGAATTAATATAAAGTGTATTCTTAA
```

2451 TCCTAGCCTTATCCCTAAATATATATATATATATTGTATACTCTAGCTAG
2501 CTCTATGTAGGGCTAGTTCTAGTACTGCCTCTAGTTAGTTAAAAGGGAAA
2551 ACCCTTAAATAAGAAGAAAAATCCCTTTATATTTTGTCAGGCGAAAACAA
2601 CCACCCGAAAACGACGGATTTGACGATGACACTAACAACAAAGCTAACGA
2651 ATTTGACGATATTAGCAATTGAACCTAGATATCGGGATCTAGGTCTGCGA
2701 GGTTTCCGATCCACGCCTAGGATTCAAGCTAGGGGGTAGGGTCTTTTTCT
2751 AATAATAGGTTATTTTATTAATTAAACAATCCAAGCCTAAGGCAACGAAG
2801 GGAGAGTAAAGTTTCTATTTAAAGGGAGGGAATCTAGGGGTTTTATCTAG
2851 CTAGGAGGTCACATGACTAGGGATCCGATGTGGCCGAATTGATCTGACAA
2901 GCCAATAGATCTGACGAAGCCAAGGTCTAGGGGCCCGAGGTCTTGTGAGA
2951 GAGGTCTCGAGAGGTCACAATGCTAGCCACACAATATCTATCAATATATG
3001 AATATATTATATTATATGATTTACCCTAGATAGCAATTTATGCCATTAAC
3051 CAGTACTCCTGCCGTGATGTTGCTTTGTAGTAGGAAAACCATACTAGGTT
3101 GCTAATTATCTAGATAACTAGATAACTAGTTAGTTGCCTAGTTAGAACTC
3151 GTATCTCAAATCCCTGTTACGTATCTCTCTACCCGCAGTCCTTTTTAGAT
3201 CTTGTTATTGAGTCTCGTAGAAGTAGCACATCCGCGCTACCTGCAGCTGG
3251 ACCAGCTATGAGACTGACAAAAAACATCCTTACCATAACTCGTAAGCTCA
3301 AGTGTTTATTTTCTGCTTCAAGTGCTTGAGAAAATAGCCCCACGGTCAAG
3351 AAAAATCCACTTGATGTACCAGTCATCTCATTAATCTGTCTGAGTCTAGC
3401 ATGTCGTGCAGCGATCTCGGAACACGGAAACTGCGAGCAATCGGGTACAC
3451 CAAGGAGGCTATTCCCTATATGAAAGGGAGCAGTGGCGTCTCTGTGAAGG
3501 AGAGTCGCCACGATCGCTACCATAAAAATGCCAATGTGGCTTATACCAGT
3551 GCACCAGAAAATAGTCCTTAGGAAAGCCTTCTCTTGCCTCCTCGGCCACG
3601 CTGTTACTAATTTCTCGGCACGATATTGATTTAGGATCCACAGTGAAAAG
3651 ACGGGAAAGGCAGTGGAAAGTCCAACTGTGTAAGAGAGATAGCCTAGTGC
3701 GGCCAAACTTCTTCAAAAAGTAAGCATAGTCAGTGAGTCAGAGTTAACAG
3751 GGAATCACATACTCAAACTTGCGGAGGAATGCGCCATGCGGTACGGTCTC
3801 ATGCAGAATTATCAAAATGAGCCCAACCAGCTGAGCAATGTAAAGCATTA
3851 GGTGAAGCCAAAACCAAGGCCCATTATCCCAAATGGACTGCATCGACGCA
3901 ACAGCGCGAAACCCGAACCATGGTGATGTGGTTCCATAGCTTAATGTAGC
3951 ATCCGAAGAATCAATGAACTGTAATGGGCAGGGAAAGTCAATGATCGGAT
4001 ATCCTTCCCGTGACTTCCATATTACGCCGGCTAAACAAAAGAAACCCTGC
4051 AGAGAGATAAAGATCCAATCACTTCGCGACATAGGGAAAAATAGAGGAAA
4101 ACTGATAATAACTTTAGGTCCAGTTTCATGCAATATTGGGAAAGGCCAGA
4151 AGCATAATCCGTACAATCGTCATGATATCGTCAAAGCGAGACTAAGCTGT
4201 TTCTTTATAGGGGCTGAGAAATCTTGGCAATAGGAAACCGGAAGAATGCC
4251 GAGTGCGACTGACGCAAAGAATTGGCTTGAGCACCCGACCCCCTCTCCAT
4301 CCCTAACCCGTGTCGTCATTATCTTTCGGCAATAGATATGGCGTTTCATT
4351 TCACTGTAACATACAGATTACTCCGTATTTATGTAATAATACACCCTATT
4401 ACATGTAATATTACACGTAGGGAGGGGGTGATTAGGAAGCGTGCGGATGA
4451 TACGTAGAACTACTATATAATTAACTACTCCGTATAGATAGCTAGTATTA
4501 GTTATTGTAAAGGTAGGGGTCAATATAGATGATTAAAAGCGTTCAATTTA
4551 GTCAATTAGAGGTGCAGACAGCACCTGAGTTTTGTACCTAAAAGGTACAT
4601 AGTGCGCTATAGTAATGACTAGTTTACGGAGGTACTTCTAATACATTGTA
4651 TCCACTCGTTGTCTTAGAGAGAGTTTTATCCTAGTCAATGCGCGCTGCCT
4701 CATACATCCTAGGCTTTAAGGGAGCTCTCCCTGACAGTTATTGCAGCTAC
4751 CTTAGCTACATTCAGGGGTGCTATTTACGCATAAGGGTGTGCTTAATAAA
4801 CACACCCCTGTCAATACCCAAGCCACAATAAAGACAGTTTTTGTCTTTGT
4851 GCAGATTCGTGAATCCTACTAAAGCTTACAGACACATGCAATACCACTAA
4901 TAAAATATTGATTTGGAGTTGTTTTGGAGGTGGATTTTAGTATAGGACTA
4951 TAACCACTCTCCTATCTTACATCAGAATAAACCCAATTTTTGTGGTCTAG
5001 ACAAAACGTAATGCTAAGCAAAAAGTGGAGAGCTTGCAAAAGCCAGAGAG

```
5051  AAGACATGGCGCCATAACTAAATTGATCCTTGTATATCTGATGCAGTTGC
5101  CACTGCGTGAGAGATAAAGCAAGTTAATCGATTAGTATCCGATCAAAACT
5151  TTTCGTTCTAGGAAAGCTTTATTTCGCACACATCAATGTTCTTGGAATGC
5201  TAACCCGAATCGCAATTATCTGAAACC[redacted]ACTAGCGACTTCAAGGTAAT
5251  AATCGTGGGAGGATCAGTGGCTGGGCTTTCACTAGCCCACTGCTTAGAAA
5301  AAATCGGTGTTTCTTTCATGGTTCTAGAGAAGGGTAATCAAATAGCTCCC
5351  CAACTCGGTGCCTCAATTGGCATTTTGCCAAATGGTGGACGTATTCTTGA
5401  TCAACTGGGCATCTTCCATAGCATCGAGGATGAAATCGAACCTCTAGAAT
5451  CTGCTATGATGAGATACCCGGATGGTTTCTCTTTCAAAAGTCAATATCCC
5501  CAAGCTTTGCATACTAG[redacted]
5551  [redacted]TTTTGGTTATCCCGTGG
5601  CTTTCCTTGAGAGGCAAAGGTTTCTTCAGATACTTTATGATAAACTCAAG
5651  AGCAAAGACTGCGTTTTTACAAACAAGCGGGTAGTCAGTATTGCAAGTGG
5701  CCAAGACAAAGTCACAGCAAAGACTTCAGATGGCGCTAAGTACTTAGCAG
5751  ATATCGTGATCGGTGCTGACGGGGTCCACAGCATCGTCAGGTCAGAGATT
5801  TGGAGGCATTTGAAGGAAAACTCTCAAATATCAGTATTAGAGGCACCGAA
5851  CGCAA[redacted]
5901  [redacted]GTATTAAGCATGATTATTCATGCATTTACGGAATTTCTT
5951  TAAACGTTCCCCAGATCATCCTAGGAATACAGTTAAACTGTTTAGATGAC
6001  GGAGTGTCAATACACTTGTTTACGGGTAAACAATCCAAATTATTTTGGTT
6051  TGTTATCATCAAAACGCCTCAGGCTAGCTTTGCTAAAGTAGAGATTGACA
6101  ATACACATACAGCAAGGTGTATCTGCGAAGGACTGAGGACGAAAAAGGTT
6151  TCAGATACCTTATGTTTTGAAGATGTATGGTCAAGATGCACCATATTCAA
6201  GATGACGCCTCTTGAGGAAGGGGTGTTTAAGCATTGGAACTATGGCCGCT
6251  TAGCATGTATTGGTGATGCTATCCGCAAG[redacted]
6301  [redacted]
6351  [redacted]ATGGCCCCAAATAAT
6401  GGGCAAGGAGCAAATATGGCGATAGAGGACGCTTGCAGTCTCGCAAACAT
6451  CCTCCAGAAAAAGATATCACATGGTTCGATTCGAGACCAAGATATCAATT
6501  CAATGTTTCAGGAATTCTCTATGGCTCAACGGGCTCGCACGGAGAGCGTC
6551  TGCGCGCAGTCGGAGTTTCTAGTCCGCATGCATGCGAATCAAGGTATTGG
6601  AAGAAGACTTCTTGGGCGGTACCTTATTCCTTTCCTGTATGACGCACCTG
6651  CTGGTTTATCTGGATTTTCTATAAGTGGCGCAACAAGAATAGAGTTCATA
6701  GACTTGCCCACTAGATCTCTTAGGGGAGCGTGGGGAAAGTCATGGAGAGG
6751  GTCATGGGAATTCATCCTACAAAGCTTGGTCTATTTGCGACCCAAGTTTA
6801  GGATAGTTTATGCCTTGTATCTCGTTGCAGCTGCAGCTTTTATCTTGTAT
6851  TGTCTTAGCAGTCTCTTCCCG[redacted]CAAGGAACAACTGTCGAAAATGGCCT
6901  TAATCTGGAAAAGCTAATGCGGCGATGAAGGCAGGCAGAACTCAAAAACA
6951  GACAAGCAATGACCCTCATATTGTTAAATGCTAGTTGTTACATAACTTCA
7001  TGTGATTCGAGGTGAAACTATATTAACCCATTTTCCAACTAGGAGAAAAA
7051  TGTGTTATAGAAAAGTAAGCAAATAGCTAGTAAGAATATAATAAAAAGCT
7101  AGACATGAACTTATATTTCCAACAGCAAGACCTAGGTATATAGTAACTAA
7151  AAGGTATTACGAACCTAACATATACTAATAGTATATAATAGAGTAGCTTA
7201  TGTAGAAATATAAGTAAAGAAATAGCAAATAGGTAAGGAATTAATAAACC
7251  TAATAGGCCATAGTAGCACCATTTAGACTAAACACAATATAGTTAGCTAT
7301  AGTTATGTAGTCATAACTAAGAATTCAATTAAGTAAACACTTAGTAAGAT
7351  AGTAATAAGTTACTATAGAGAATATAGAGTCTATATCCTTATCCTTGTTC
7401  ATAGTGTCTATAAGCTCCTAGAGCTATTCTAGAATAGCAAAACGATTAGC
7451  AAAATTGCCCTCAAGTGTAAGAATAGCCTAGTGTAAAAACCATAGCGTTA
7501  AGAAACTATAAGACTAGTAAAAAAAAGGGAGACTTGTAGTCTTGCAGGTA
7551  TTGCCTCTCTTATTACACTAGATATAGCGCTTTAAAGTTTAGTCTTAGCT
7601  AGAGTAGAAATTAAAACCTAATGGAAACTCAAGTTGATTTATAGTAATAT
```

```
7651   AGCCTTAATAAGGGGTTTTTTTTAAAGTCCGTGTACTTAGTATGTAAATA
7701   ACACATATAGCTACACTTTTCAAAGGAAATTGTAGTTATATTAGTGGTAA
7751   AACGGTGGTAAATAGAAGGGTTAAAGAGGGTATGAACTAAGCTTAAAAAA
7801   ACCCTAGGAAAGAAACTAGGTTTATAGGGAGAAAAACCTAATCAGGCAAT
7851   AGGGAACTGCAAGTAAATGTTAGAGATAGGATACTTACAAAATAAAGGGC
7901   TAGGAAAACTTTAGATCCTTTAGATAATTAAGCAGCTAGCTAGCTATGGG
7951   ATAGCTATGTGTTTATAAAGCAAGGTATTTAGCAAAGACTACTTATACTA
8001   TATATAGTAAATTAGAGTTTAAGACCTTTACACACCTACTCCTAGGTAGT
8051   ATCTTTCTAGTAGTAACTACGAATCTTAGCCTTCAATCTATTCATTACCC
8101   TATAACCGAAGTTATAACAAATCCTTAAATTTTTAATAAGTATTAATCTA
8151   TACTTAACACATATAAGTACTATATTTATCAAGTATTAATTAACACTATA
8201   AAGGTTATAAATATAAATTCTACTTATAAAAAGGAAATATATCTTCTTTA
8251   AAATAAGGGCTAATTAATTAATTTAATGACGCATGAAAATATTATTGTTA
8301   TAAAGGAAAAGGGGGGATTATTTACTACCCCTTAAGTTATATAATCATGC
8351   GTTGTTAGAAATATTAAAGCTTCTAGTGTAAAATAAAAGCTAAGTGCAAC
8401   TAAGTGTAATTAAAAGCACTAGGCTTATAACCTATAAGATAGTGGAAAAA
8451   GTAATAATAATAAATTCAGCTATCTAAGCTCTTTATATACGTGGTATAAT
8501   AAGGCTATATAACGAGAGCAAAAGACAGTCTTTACCCTAAGTGACAAGGT
8551   CTCGTAATTAGCCGCGAAGAGGGAAAGCATCGCGATGAAAGTGATGCCTA
8601   AGATGTGAGGCTGCTACATCTAACAGATCAGACCCTTCGTCTCCTCAGAA
8651   CACGCGGTTTGAAAAGTTCTACCTCTAGCAACTCCTCGCACCAAGCTGTT
8701   TCTACATGCTCTTACCGCAATCTAAACTGAAACCCAAAATTCACCTCGCA
8751   CATAGCCCCTAATCCGCAATTGCTTTAAC[redacted]CAATACGGTAATTTAACA
8801   ACTGTATTACTTCTGCGTAATACTTTATTGTCCTTGAATTCTTCGTCAAT
8851   CTGCCATGTTCACTGGCTGCAAGTGATTGTGGCTCTGCTTGTCTTGATCG
8901   TCTGCATCTTTCTATATTGGCGAACACCCACTGGCATCAATGCTCCTTTC
8951   GCAGGATATCGTTCACCATGGGAGCCGCCGCTCTTGGTTCAGATGCGTTA
9001   CGTCTTCAACGCTGCCTCAATGATACGCGAAGGATATGCTAAG[redacted]
9051   [redacted]
9101   [redacted]TGGAAAGACTCCTTGTTCCAGATCTCACGATACGACGGTGA
9151   CATTCTTATTGTGCCTCCAAGATATTTGGATGACCTCCACAACAAGTCAC
9201   AAGAGGAGTTAAGTGCTATTTATGGTTTGATTCGG[redacted]
9251   [redacted]
9301   [redacted]AATTTTGGTGGTAGCTATAGCGGCATCACCCTGCTTGGAGAAAAC
9351   GATGTTGGCATTCGTGCGCTTCAG[redacted]
9401   [redacted]ACAAAAATCACCCCAAATCTTGCG
9451   AAATTATGCGATGACATAAGGGATGAGTTTCAGTATTGTCTAGATACAGA
9501   CTTCCCAGCCTGCAGAG[redacted]
9551   [redacted]ATTGGACATCAGTGTC
9601   CGTGCATCCATTGTTTCTAAAAGCAGTCGAAAGGATAACACATCGGATTT
9651   TTGTTGGATTGCCATTATGTCGGAATCCCCAATGGGTCCAAGCGACCAGC
9701   AAGCATGCACATTACG[redacted]
9751   [redacted]CAACAATGATACAGATAGCTATGAGATCTGTCC
9801   CAAAGTTCATTCAGCCTTTACTAAATTTTTGCCTTCCGTGGCCATGGAAG
9851   AACGCAGCCTGTGTTCGTGAAGCAAAGAATGCCCTTATATTAGAAATGCA
9901   ACGCCGACGAAATCTCGAGAAAGTTAACAGTTTTGATTATATCAAATCCA
9951   ATGACTTGCTGCAAGCAGTTATGGAAATGTCTTCTCCTAGTCATGAGGAT
10001  AGCCAGCTTGATGTTGTCGCCCAGATAATGCTCACGATGAACACAATCGC
10051  TGGCCACAGTACTGCCGCATCCGGAGCACATGCACTGTTCGATATGGTTA
10101  GCCACTCTAAGTATATTGAATTGCTGCGTGAGGAGGCTCTTCAAGTCTTT
10151  CGACATGTTGAACTGCGTGTTACAAAACAGGCTTTGGGGGATTTGCGAAA
10201  ATTGGACAGCTTCCTCAGAGA[redacted]
```

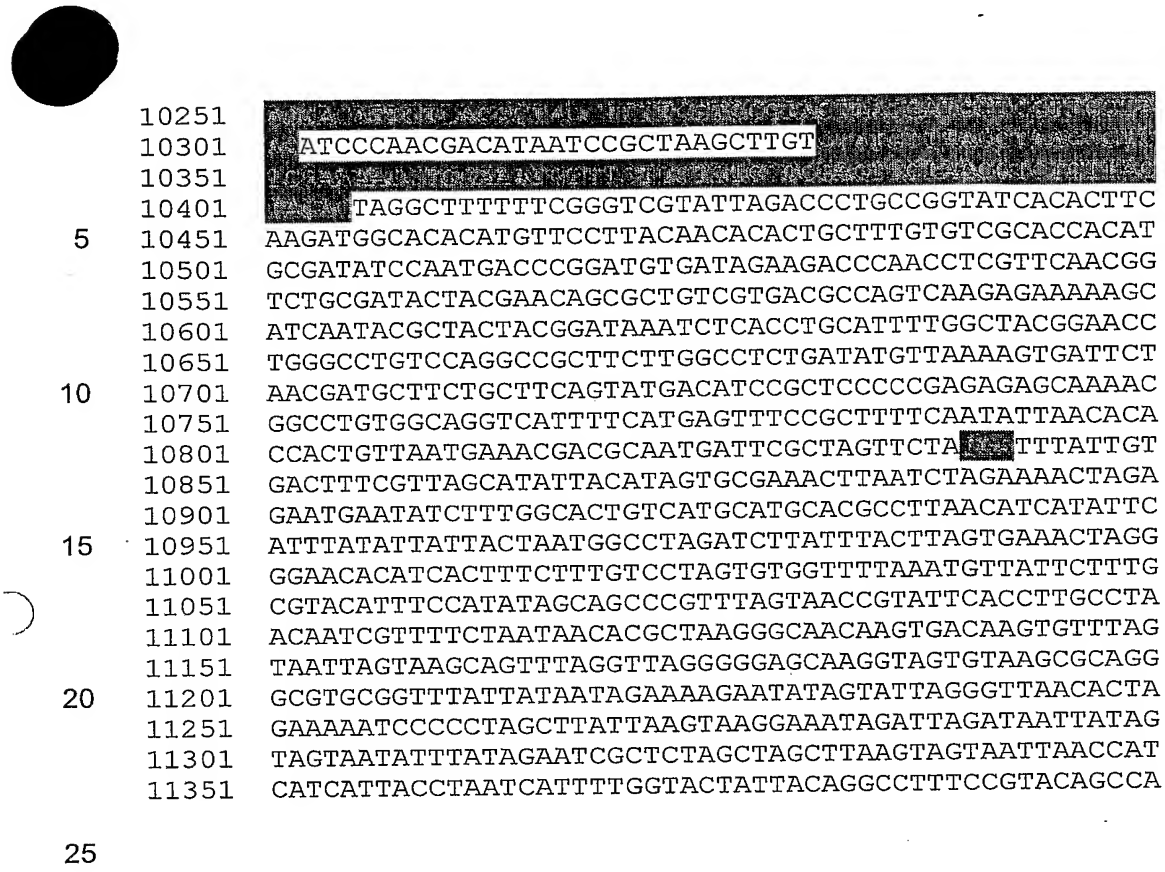

10251
10301 ATCCCAACGACATAATCCGCTAAGCTTGT
10351
10401 TAGGCTTTTTTCGGGTCGTATTAGACCCTGCCGGTATCACACTTC
10451 AAGATGGCACACATGTTCCTTACAACACACTGCTTTGTGTCGCACCACAT
10501 GCGATATCCAATGACCCGGATGTGATAGAAGACCCAACCTCGTTCAACGG
10551 TCTGCGATACTACGAACAGCGCTGTCGTGACGCCAGTCAAGAGAAAAAGC
10601 ATCAATACGCTACTACGGATAAATCTCACCTGCATTTTGGCTACGGAACC
10651 TGGGCCTGTCCAGGCCGCTTCTTGGCCTCTGATATGTTAAAAGTGATTCT
10701 AACGATGCTTCTGCTTCAGTATGACATCCGCTCCCCCGAGAGAGCAAAAC
10751 GGCCTGTGGCAGGTCATTTTCATGAGTTTCCGCTTTTCAATATTAACACA
10801 CCACTGTTAATGAAACGACGCAATGATTCGCTAGTTCTA TTTATTGT
10851 GACTTTCGTTAGCATATTACATAGTGCGAAACTTAATCTAGAAAACTAGA
10901 GAATGAATATCTTTGGCACTGTCATGCATGCACGCCTTAACATCATATTC
10951 ATTTATATTATTACTAATGGCCTAGATCTTATTTACTTAGTGAAACTAGG
11001 GGAACACATCACTTTCTTTGTCCTAGTGTGGTTTTAAATGTTATTCTTTG
11051 CGTACATTTCCATATAGCAGCCCGTTTAGTAACCGTATTCACCTTGCCTA
11101 ACAATCGTTTTCTAATAACACGCTAAGGGCAACAAGTGACAAGTGTTTAG
11151 TAATTAGTAAGCAGTTTAGGTTAGGGGGAGCAAGGTAGTGTAAGCGCAGG
11201 GCGTGCGGTTTATTATAATAGAAAAGAATATAGTATTAGGGTTAACACTA
11251 GAAAAATCCCCCTAGCTTATTAAGTAAGGAAATAGATTAGATAATTATAG
11301 TAGTAATATTTATAGAATCGCTCTAGCTAGCTTAAGTAGTAATTAACCAT
11351 CATCATTACCTAATCATTTTGGTACTATTACAGGCCTTTCCGTACAGCCA

Figure 11.

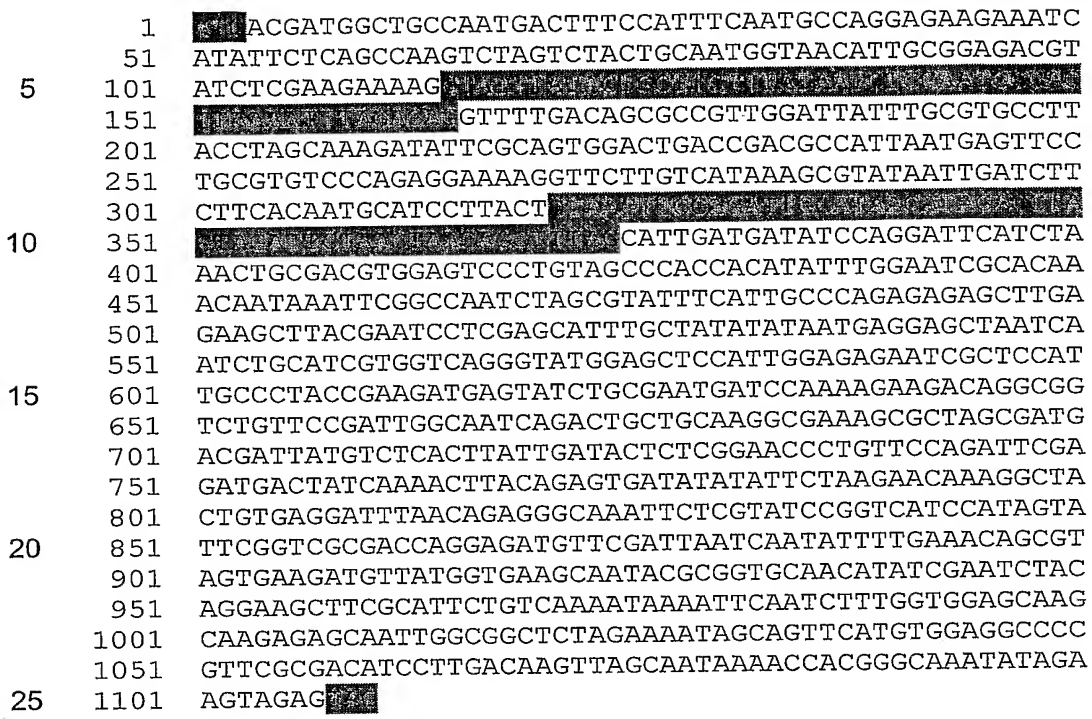

1 ACGATGGCTGCCAATGACTTTCCATTTCAATGCCAGGAGAAGAAATC
51 ATATTCTCAGCCAAGTCTAGTCTACTGCAATGGTAACATTGCGGAGACGT
101 ATCTCGAAGAAAAG
151 GTTTTGACAGCGCCGTTGGATTATTTGCGTGCCTT
201 ACCTAGCAAAGATATTCGCAGTGGACTGACCGACGCCATTAATGAGTTCC
251 TGCGTGTCCCAGAGGAAAAGGTTCTTGTCATAAAGCGTATAATTGATCTT
301 CTTCACAATGCATCCTTACT
351 CATTGATGATATCCAGGATTCATCTA
401 AACTGCGACGTGGAGTCCCTGTAGCCCACCACATATTTGGAATCGCACAA
451 ACAATAAATTCGGCCAATCTAGCGTATTTCATTGCCCAGAGAGAGCTTGA
501 GAAGCTTACGAATCCTCGAGCATTTGCTATATATAATGAGGAGCTAATCA
551 ATCTGCATCGTGGTCAGGGTATGGAGCTCCATTGGAGAGAATCGCTCCAT
601 TGCCCTACCGAAGATGAGTATCTGCGAATGATCCAAAAGAAGACAGGCGG
651 TCTGTTCCGATTGGCAATCAGACTGCTGCAAGGCGAAAGCGCTAGCGATG
701 ACGATTATGTCTCACTTATTGATACTCTCGGAACCCTGTTCCAGATTCGA
751 GATGACTATCAAAACTTACAGAGTGATATATATTCTAAGAACAAAGGCTA
801 CTGTGAGGATTTAACAGAGGGCAAATTCTCGTATCCGGTCATCCATAGTA
851 TTCGGTCGCGACCAGGAGATGTTCGATTAATCAATATTTTGAAACAGCGT
901 AGTGAAGATGTTATGGTGAAGCAATACGCGGTGCAACATATCGAATCTAC
951 AGGAAGCTTCGCATTCTGTCAAAATAAAATTCAATCTTTGGTGGAGCAAG
1001 CAAGAGAGCAATTGGCGGCTCTAGAAAATAGCAGTTCATGTGGAGGCCCC
1051 GTTCGCGACATCCTTGACAAGTTAGCAATAAAACCACGGGCAAATATAGA
1101 AGTAGAG

Figure 12.

```
      1  [redacted]ACTAGCGACTTCAAGGTAATAATCGTGGGAGGATCAGTGGCTGGGCT
     51  TTCACTAGCCCACTGCTTAGAAAAAATCGGTGTTTCTTTCGTGGTTCTAG
    101  AGAAGGGTAATCAAATAGCTCCCCAACTCGGTGCCTCAATTGGCATTTTG
    151  CCAAATGGTGGACGTATTCTTGATCAACTGGGCATCTTCCATAGCATCGA
    201  GGATGAAATCGAACCTCTAGAATCTGCTATGATGAGATACCCGGATGGCT
    251  TCTCTTTCAAAAGTCAATATCCCCAAGCTTTGCATACTAG[redacted]
    301  [redacted]
    351  [redacted]TTTTGGTTATCCCGTGGCTTTCCTTGAGAGGCAAAGGTTTCTTC
    401  AGATACTTTATGATAAACTCAAGAGCAAAGACTGCGTTTTTACAAACAAG
    451  CGGGTAGTCAGTATTGCAAGTGGCCAAGACAAAGTCACAGCAAAGACTTC
    501  AGATGGCGCTAAGTACTTAGCAGATATCGTGATCGGTGCTGACGGGGTCC
    551  ACAGCATCGTCAGGTCAGAGATTTGGAGGCATTTGAAGGAAAACTCTCAA
    601  ATATCAGTATTAGAGGCACCGAACGCAA[redacted]
    651  [redacted]GTATTAAGCATGATTA
    701  TTCATGCATTTACGGAATTTCTTTAAACGTTCCCCAGATCATCCTAGGAA
    751  TACAGTTAAACTGTTTAGATGACGGAGTGTCAATACACTTGTTTACGGGT
    801  AAACAATCCAAATTATTTTGGTTTGTTATCATCAAAACGCCTCAGGCTAG
    851  CTTTGCTAAAGTAGAGATTGACAATACACATACAGCAAGGTGTATCTGCG
    901  AAGGACTGAGGACGAAAAAGGTTTCAGATACCTTATGTTTTGAAGATGTA
    951  TGGTCAAGATGCACCATATTCAAGATGACGCCTCTTGAGGAAGGGGTGTT
   1001  TAAGCATTGGAACTATGGCCGCTTAGCATGTATTGGTGATGCTATCCGCA
   1051  AG[redacted]
   1101  [redacted]
   1151  [redacted]ATGGCCCCAAATAATGGGCAAGGAGCAAATATGGCGATAGAG
   1201  GACGCTTGCAGTCTCGCAAACATCCTCCAGAAAAAGATATCACATGGTTC
   1251  GATTCGAGACCAAGATATCAATTCAATGTTTCAGGAATTCTCTATGGCTC
   1301  AACGGGCTCGCACGGAGAGCGTCTGCGCGCAGTCGGAGTTTCTAGTCCGC
   1351  ATGCATGCGAATCAAGGTATTGGAAGAAGACTTCTTGGGCGGTACCTTAT
   1401  TCCTTTCCTGTATGACGCACCTGCTGGTTTATCTGGATTTTCTATAAGTG
   1451  GCGCAACAAGAATAGAGTTCATAGACTTGCCCACTAGATCTCTTAGGGGA
   1501  GCGTGGGGAAAGTCATGGAGAGGGTCATGGGAATTCATCCTACAAAGCTT
   1551  GGTCTATTTGCGACCCAAGTTTAGGATAGTTTATGCCTTGTATCTCGTTG
   1601  CAGCTGCAGCTTTTATCTTGTATTGTCTTAGCAGTCTCTTCCCG[redacted]
```

Figure 13.

```
   1  NNNCAATACGGTAATTTAACAACTGTATTACTTCTGCGTAATACTTTATT
  51  GTCCTTGAATTCTTCGTCAATCTGCCATGTTCACTGGCTGCAAGTGATTG
 101  TGGCTCTGCTTGTCTTGATCGTCTGCATCTTTCTATATTGGCGAACACCC
 151  ACTGGCATCAATGCTCCTTTCGCAGGATATCGTTCACCATGGGAGCCGCC
 201  GCTCTTGGTTCAGATGCGTTACGTCTTCAACGCTGCCTCAATGATACGCG
 251  AAGGATATGCTAAG[redacted]
 301  [redacted]TGGAAAGACTCCTTGTTCCA
 351  GATCTCACGATACGACGGTGACATTCTTATTGTGCCTCCAAGATATTTGG
 401  ATGACCTCCACAACAAGTCACAAGAGGAGTTAAGTGCTATTTATGGTTTG
 451  ATTCGG[redacted]
 501  [redacted]AATTTTGGTGGTAGCTATAGCGGC
 551  ATCACCCTGCTTGGAGAAAACGATGTTGGCATTCGTGCGCTTCAG[redacted]
 601  [redacted]ACA
 651  AAAATCACCCCAAATCTTGCGAAATTATGCGATGACATAAGGGATGAGTT
 701  TCAGTATTGTCTAGATACAGACTTCCCAGCCTGCAGAG[redacted]
 751  [redacted]
 801  [redacted]ATTGGACATCAGTGTCCGTGCATCCATTGTTTCTAAAAGCAGTCG
 851  AAAGGATAACACATCGGATTTTTGTTGGATTGCCATTATGTCGGAATCCC
 901  CAATGGGTCCAAGCGACCAGCAAGCATGCACATTACG[redacted]
 951  [redacted]CAACAATGATAC
1001  AGATAGCTATGAGATCTGTCCCAAAGTTCATTCAGCCTTTACTAAATTTT
1051  TGCCTTCCGTGGCCATGGAAGAACGCAGCCTGTGTTCGTGAAGCAAAGAA
1101  TGCCCTTATATTAGAAATGCAACGCCGACGAAATCTCGAGAAAGTTAACA
1151  GTTTTGATTATATCAAATCCAATGACTTGCTGCAAGCAGTTATGGAAATG
1201  TCTTCTCCTAGTCATGAGGATAGCCAGCTTGATGTTGTCGCCCAGATAAT
1251  GCTCACGATGAACACAATCGCTGGCCACAGTACTGCCGCATCCGGAGCAC
1301  ATGCACTGTTCGATATGGTTAGCCACTCTAAGTATATTGAATTGCTGCGT
1351  GAGGAGGCTCTTCAAGTCTTTCGACATGTTGAACTGCGTGTTACAAAACA
1401  GGCTTTGGGGGGATTTGCGAAAATTGGACAGCTTCCTCAGAGA[redacted]
1451  [redacted]
1501  [redacted]ATCCCAACGACATAATCCGCTAAGCTT
1551  GT[redacted]
1601  [redacted]TAGGCTTTTTTCGGGTCGTATTAG
1651  ACCCTGCCGGTATCACACTTCAAGATGGCACACATGTTCCTTACAACACA
1701  CTGCTTTGTGTCGCACCACATGCGATATCCAATGACCCGGATGTGATAGA
1751  AGACCCAACCTCGTTCAACGGTCTGCGATACTACGAACAGCGCTGTCGTG
1801  ACGCCAGTCAAGAGAAAAAGCATCAATACGCTACTACGGATAAATCTCAC
1851  CTGCATTTTGGCTACGGAACCTGGGCCTGTCCAGGCCGCTTCTTGGCCTC
1901  TGATATGTTAAAAGTGATTCTAACGATGCTTCTGCTTCAGTATGACATCC
1951  GCTCCCCCGAGAGAGCAAAACGGCCTGTGGCAGGTCATTTTCATGAGTTT
2001  CCGCTTTTCAATATTAACACACCACTGTTAATGAAACGACGCAATGATTC
2051  GCTAGTTCTA[redacted]
```

Figure 14.

```
  1  MTMAANDFPF QCQEKKSYSQ PSLVYCNGNI AETYLEEKVL TAPLDYLRAL
 51  PSKDIRSGLT DAINEFLRVP EEKVLVIKRI IDLLHNASLL IDDIQDSSKL
101  RRGVPVAHHI FGIAQTINSA NLAYFIAQRE LEKLTNPRAF AIYNEELINL
151  HRGQGMELHW RESLHCPTED EYLRMIQKKT GGLFRLAIRL LQGESASDDD
201  YVSLIDTLGT LFQIRDDYQN LQSDIYSKNK GYCEDLTEGK FSYPVIHSIR
251  SRPGDVRLIN ILKQRSEDVM VKQYAVQHIE STGSFAFCQN KIQSLVEQAR
301  EQLAALENSS SCGGPVRDIL DKLAIKPRAN IEVE
```

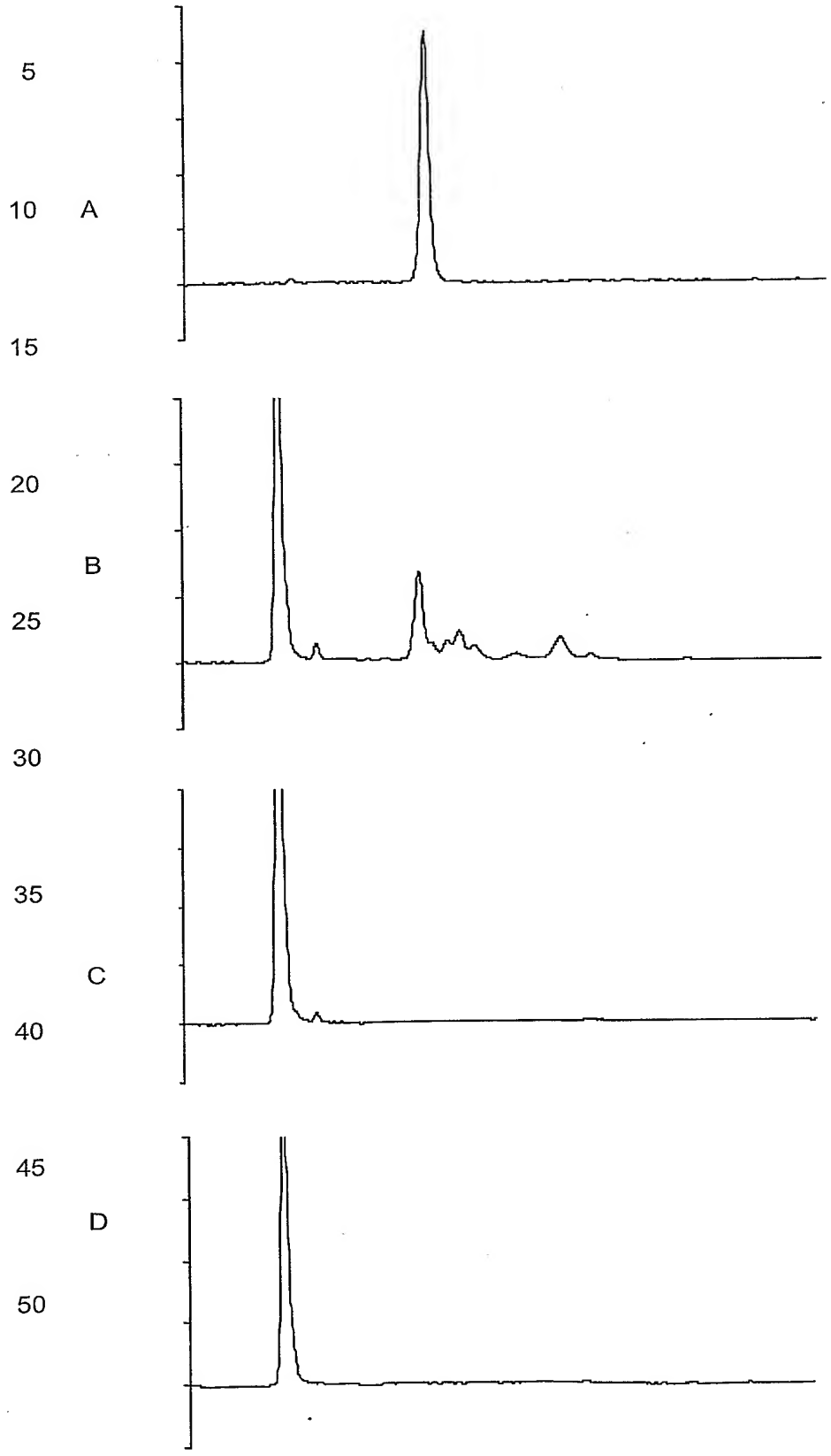
Figure 15.

```
  1  MTSDFKVIIV GGSVAGLSLA HCLEKIGVSF VVLEKGNQIA PQLGASIGIL
 51  PNGGRILDQL GIFHSIEDEI EPLESAMMRY PDGFSFKSQY PQALHTSFGY
101  PVAFLERQRF LQILYDKLKS KDCVFTNKRV VSIASGQDKV TAKTSDGAKY
151  LADIVIGADG VHSIVRSEIW RHLKENSQIS VLEAPNASIK HDYSCIYGIS
201  LNVPQIILGI QLNCLDDGVS IHLFTGKQSK LFWFVIIKTP QASFAKVEID
251  NTHTARCICE GLRTKKVSDT LCFEDVWSRC TIFKMTPLEE GVFKHWNYGR
301  LACIGDAIRK MAPNNGQGAN MAIEDACSLA NILQKKISHG SIRDQDINSM
351  FQEFSMAQRA RTESVCAQSE FLVRMHANQG IGRRLLGRYL IPFLYDAPAG
401  LSGFSISGAT RIEFIDLPTR SLRGAWGKSW RGSWEFILQS LVYLRPKFRI
451  VYALYLVAAA AFILYCLSSL FP
```

Figure 16.

```
  1  MQYGNLTTVL LLRNTLLSLN SSSICHVHWL QVIVALLVLI VCIFLYWRTP
 51  TGINAPFAGY RSPWEPPLLV QMRYVFNAAS MIREGYAKWK DSLFQISRYD
101  GDILIVPPRY LDDLHNKSQE ELSAIYGLIR NFGGSYSGIT LLGENDVGIR
151  ALQTKITPNL AKLCDDIRDE FQYCLDTDFP ACRDWTSVSV HPLFLKAVER
201  ITHRIFVGLP LCRNPQWVQA TSKHAHYATM IQIAMRSVPK FIQPLLNFCL
251  PWPWKNAACV REAKNALILE MQRRRNLEKV NSFDYIKSND LLQAVMEMSS
301  PSHEDSQLDV VAQIMLTMNT IAGHSTAASG AHALFDMVSH SKYIELLREE
351  ALQVFRHVEL RVTKQALGDL RKLDSFLRES QRHNPLSLLG FFRVVLDPAG
401  ITLQDGTHVP YNTLLCVAPH AISNDPDVIE DPTSFNGLRY YEQRCRDASQ
451  EKKHQYATTD KSHLHFGYGT WACPGRFLAS DMLKVILTML LLQYDIRSPE
501  RAKRPVAGHF HEFPLFNINT PLLMKRRNDS LVL
```

Figure 17



James & Wells Ref: 42616/29

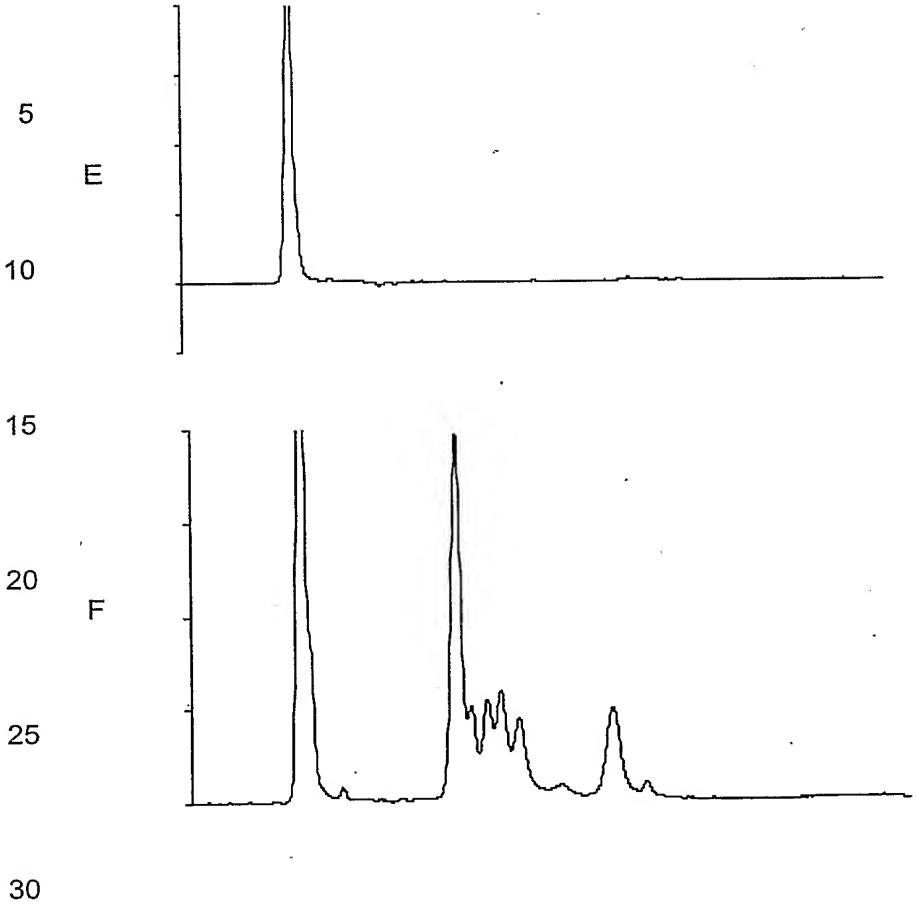
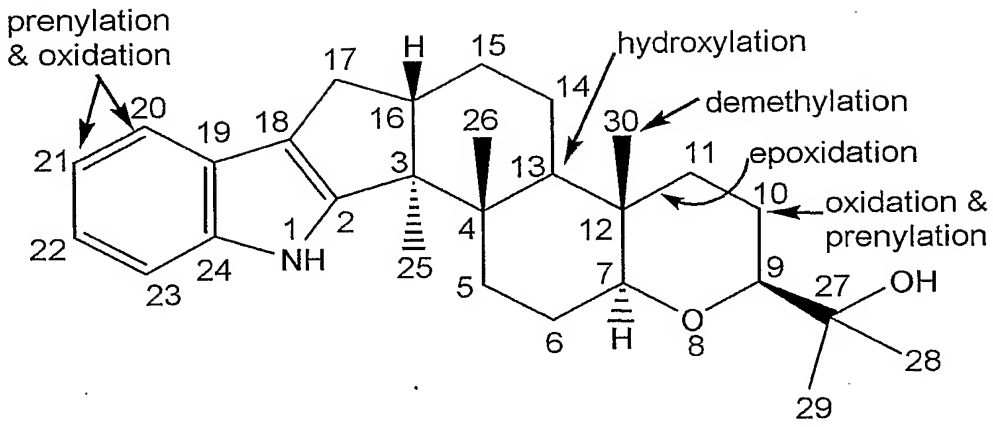

5
E
10
15
20
F
25
30

Figure 18.



Paspaline

Figure 19.

```
  1  TACTGTCCGTATTGATACACTTCGCTGGAATCGCAATCTGCTGTCCCGGC
 51..TGTAGCTGGTAGCCATCGGACAATGTAATTCGTTCTCGGACAATGCGTCT
101..AAAAGTGAACTCGCCCACTGGATTGATTCGTTGTGACTCTCTCATCAGGC
151  TATCCAATTGTTTCATATCAAGCATAGTCTGTTGCGTGATTTCCGTCCAA
201  CCATGCTTGAGCACGTTTTGCTGTTCTCTTCGGAGGCTCTCCTTCAAGCT
251  GTCGTCCGAGATCATGTCAAAAAGTATGTGAGTTAACGCCATAGCTGTTG
301  TATGAATGACAGCCATGCTCAATATTCCTAGCGTA
```

Figure 20.

```
  1  GCCCTTAGCGTGGTCGCGGCCGAGGTACCAAACGGAAAGAATGTATACCA
 51  ACCATTCGACGTTGGCCTCGATTACTCCCCCAATTCCTCGATCGGCTGAG
101  TTATAATGACCATGCCGCCCGCCTAGTCAAACATGGCTATGAGAAGCACA
151  AAAATCAACCGTTTAGGCTACTTAAGATGGACATGGATCTGATTGTCATT
201  CCTTTACAATACGCGCTGGAATTACGGGCGGTTACGAGCGACAAATTAGA
251  CCCTTTAACAGCCAGCTTTGATGACAATGCTGGTAAAGTTACGAGGATAT
301  TATTAGGGAGCGAACTTCACACACGTGCCATACAGCAGCGTTTGACCCCA
351  AAGCTTCCACAAACTCTTCCAGTGCTATTGGATGAGCTCAATCATGCCTT
401  TGGGCAAGTCTTACCTGCCGGCAACGACGGTTCCAATGCTTGGATTTCTG
451  TCAATCCATACGAATTGGTTCTCAATCTAGCTACCCGNGCTACAGCGAGG
501  CTATTCGTTGGAGACCTGATTTGTCGAAACGAANTTTTCTCGAGACTACT
551  GCTTTCTTTTAGGCGCAACACGTTGGATACGATATCCNCCTCCCGGAGTT
601  TTGGCAATTNGTNCCCANATTATTTCGCCNGGGGGATTT
```

Figure 21.

```
  1  ACAGGAAGGACCTCGGGGAGNCCCAAGAAAAACGAAGCTCCCAAGCATCG
 51  ATTTGTCACCCCCGACAGCAACTACTTGACCTTTGGGNCCGGTAAATACG
101  TCTGCCCCGGCCGATTTATAGCGGANCACATGTTGAAGCTGATGATGACC
151  GCCGTGCTCCTGCGCTACGAGTTCAGGNGGCCTCNGGGAGTCCCTGTGCC
201  CGAAAANAGTATCGGCATGTCTTTGCTTATCCAGGCAAGCCACACTGTTG
251  ATNAACGACGCAAAGATGGCGATCAGATTCTTTAAAGTATCATTATCTGA
301  AANGAAGAAAAGAGGATGTNTNCCTCTTCCCGTNAAAACTGCTGAGTGCA
351  AGTTTGTGAAAGGAGAGGNGTTACGAACAGAATGTACCTGCCCNGGGNGG
401  CNGCTCAAGGGG
```

Figure 22.

```
  1  TCCTTTGGCAGTCCAAGTTGCTAAGGATGTAGTGGCTTCTCTGTCTGCTA
 51  CTTTTCGCCTTTCAACAAAATGGAGCGAAACTCTACTGTCCAATTTTGCA
101  GTAACACCAGACCAAGCTCGACAAGTTATTAACATGCTGCCCGAGTGGAT
151  TCAAGGCTTCGTACCTGAGGGAATGGAGTGCGATTTTCCAAAGAGAATCC
201  CGTTCGCCATGACATCATTCGACCTAAATGGCTCCAATGTAGCTATGAAG
251  CTCTACGTTAATCCAAGGGTAAAGGAGATTTTAACTGGTACTCCCTCATC
301  AGACTTGGTCTGGGAGTTCCTCCGAAATTTAACACCAGAAATGAAACCAC
351  GAGCGGTCGACTTGCTTGAGAGGTTTATTACCGATAATTCAGGCCCGTTT
```

```
401  GCTATTGAGCTTGTAGGTATTGACTGCGTTGACGACGCTCACCTATCAAA
451  TGCAAGGGTCAAGCTTTACGTTCATACCATGAGCAGCTCATTTAACACCG
501  TAAAGAATTATGTTACTCTTGGGGGTGCAATCTGGGATGAACAAACCCAA
551  AAGGGCTTAGGAATACTACAAAGTATTTGGCACCTATTGCTTCAGGAGCC
601  AGAGGGTATTTCTGACAATGGATTCGACAAGCCTGTGAACGACTCTTCCA
651  TGTTATGCCAAAAGCTATATTTTAGTTTCGAGCTACGCCCAGGTACAGAC
701  TTCCCTCAGGTGAAGAGTCGATTT
```

Figure 23.

```
  1  GGNNNANAAANAACNTCNGGNNGGGCGAATTNNNNTTCCTNGGNGNGGGG
.51  GGNNAGNGGCCGCCAGTTTTCTGGGANATCTGCAGANTTCGCCCTTTCGA
101..GNNTCCNCGCCGAAGCTCTCCCTCACTTGCANTTGCACGGGGTACTTCCT
151  CTGCANNTTCCNCACCATCANAAGNCNCNACGNCTGCTGCATACTTNANT
201  TATACTAGGTTNGTTANCCGATCATNCATGTCCNGNNGCTATTGAGCTTG
251  TAGGTCATGGACTGCTANGACGACCTNNCCTATCANATAAAAGGGCAAGC
301  TTTACGTTCATACCATGAGCAGNTCATTTNACACCGTAAAGAATATGTAA
351  CTCTTGGGGGGGCATCTGGGATGACAANCCNAAAGGCTTAGGATACTNNA
401  AGATTTGCGCCTATGCTCAGGGCANAGGGATTCTGCATGATCGNAAGCTG
451  GACANTNTCATTTTCAANGTNNTTAGTCGNCTCCCAGTCTCCGNGCGTNA
501  NGNATCACNTGNGCGTNTGGGTCACNGACANT
```

Figure 24.

```
  1  CCCTCTGGCTCCTGAAGCAATAGGNGCCAAATACTTTGTAGTATTCCTAA
.51  GCCCTTTTGGGTTTGTTCATCCCAGATTGCACCCCCAAGAGTAACATAAT
101  TCTTTACGGGGTTAAATGAGCTGCTCATGGTATGAACGTAAAGCTTGACC
151  CTTGCATTTGATAGGNGAGCGTCGTCAACGCAGTCAATACCTACAAGCTC
201  AATAGCTGACGGGCNTGATTATCGGAATAACCTCTCAGCAGGCGACCGCT
251  CGGGGTTCATTCTGGGTTAAATNCGGGNACTCCAACAAGCTGATGNGGAN
301  NCTCGCCNCCCCNTAGGNAATCANNTGGGGCGTTTAGGACGNCNGACAGT
351  GGN
```

Figure 25.

```
  1  GGACTCTCTGGCAAAGCCCGTTCATTCTCTCAACATGGAGTTCCATCCGT
 51  TGGTCGAGCAGTTAAAACAAACATTCCGTGCCTCGCCAGTCCTTTTTCTT
151  GGACGCGGTTTGCTCATCCTCGTGGTCTTCTTGATTGTCATCAACATCAT
201  CCGCCAGCAGCTCCCTCGAAGTAAATCAGAGCCGCCTTTGGTGTTTCACT
251  GGATACCGTTCATCGGCAATGCCGTTTCCTACGGTCTGGATCCATTTGTC
301  TTCTACTCGCAATGCCAGAAAAAGCATGGCGACATCTTCACTTTTATCCT
351  TTTCGGCCGAAAAATGACTGTCTACCTGGGCCTTGAAGGAAACGACTTCA
401  TTCTCAATGGCAAACTTCAAGACGTCAACGCCGAGGAGATATACGCTCCA
451  CTTACGACTCCTGTCTTCGGAAGCGACATTATCTACGACTGCCCAAACGC
501  AAAATTAATGGAGCAGAAGAAATTCGTCAAATTCGGCCTGACGCACAATG
551  CTCTGTGCTCCTATGTACCTCTCATCGAGAAGGAGGTTATTGATTACCTG
601  AAAGTGGCACCTGCATTTAAAGGCCACTCTGGTGTCGTCAACATTCCTGC
651  TGCCATGGCTGAAATCACAATCTTTACAGCGAGCAGAACGCTACAGGGCA
```

```
701  AAGAAGTCCGAAACAAGCTATCGGCTGAATTTGCAGAACTATATCACGAT
751  CTCGACCTTGGCTTCCGTCCCATCAACTTCCTCATGCCATGGGCGCCTTT
801  GCCGCAAAAATAGACGCCGAGACGCCGCCCATGCAAAGATGAGATCAATT
851  TACATCGATATTATCAACGAGCGCCGAGCGTCTGGGAAA
```

5